

# SCORE Search Results Details for Application 10

## 20081104\_154455\_us-10-578-1

[Score Home Page](#) [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ](#) [Comments / Sugg](#)

This page gives you Search Results detail for the Application 10578781 and Search Result 2008110

GenCore version 6.3  
Copyright (c) 1993 - 2008 Bioccelerat ion Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 2008, 17:10:38 ; Search time 1527 Seconds  
(without alignments)  
45700.779 Million cell updates/sec

Title: US-10-578-781-1

Perfect score: 756

Sequence: 1 at gggt gccgat at caaaaa..... aggagcaaattt gaacat tag 756

Scoring table: IDENTI TY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11299264 seqs, 46154159321 residues

Total number of hits satisfying chosen parameters: 22598528

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl : \*

1: gb\_env: \*  
2: gb\_pat: \*  
3: gb\_ph: \*  
4: gb\_pl: \*  
5: gb\_pr: \*  
6: gb\_ro: \*  
7: gb\_sts: \*  
8: gb\_sy: \*  
9: gb\_un: \*  
10: gb\_vi: \*  
11: gb\_ov: \*  
12: gb\_in: \*  
13: gb\_om: \*  
14: gb\_ba: \*  
15: gb\_htg1: \*  
16: gb\_htg2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
------------	-------	-------------	--------	----	----	-------------

1	756	100.0	756	2	DD160143	DD160143	Novel	Br e
2	756	100.0	756	2	DI 115290	DI 115290	Novel	Br e
3	366.2	48.4	1746	14	BACSP01	L47358	Paeni bacilli	
4	363.2	48.0	1799	14	Z54161	Z54161	B. coagulans	
5	361.6	47.8	3701	14	AJ704768	AJ704768	Bacillus	
c	6	361.6	47.8	110000	14	CP000560_22	Cont i nuat i on	( 23 o
c	7	354.6	46.9	110000	14	CP000813_21	Cont i nuat i on	( 22 o
c	8	353.6	46.8	110000	14	CP000557_23	Cont i nuat i on	( 24 o
c	9	348.4	46.1	110000	14	AP006627_19	Cont i nuat i on	( 20 o
c	10	345.6	45.7	864	14	BACSP02A	M15744	B. subt illis
c	11	345.6	45.7	204537	14	Z99116	Z99116	Bacillus su
c	12	345.6	45.7	282700	14	BACJH642	D84432	Bacillus su
c	13	343	45.4	110000	14	BA000004_16	Cont i nuat i on	( 17 o
c	14	342	45.2	110000	14	CP000764_28	Cont i nuat i on	( 29 o
c	15	339.6	44.9	1920	14	X63757	X63757	B. megateriu
c	16	338.8	44.8	110000	14	AE016877_40	Cont i nuat i on	( 41 o
c	17	338.2	44.7	110000	14	AE017194_38	Cont i nuat i on	( 39 o
c	18	333.8	44.2	110000	14	BA000043_23	Cont i nuat i on	( 24 o
c	19	333.4	44.1	110000	14	CP000903_39	Cont i nuat i on	( 40 o
c	20	330.2	43.7	110000	14	CP000001_39	Cont i nuat i on	( 40 o
c	21	330.2	43.7	110000	14	CP000485_39	Cont i nuat i on	( 40 o
c	22	330.2	43.7	110000	14	AE016879_39	Cont i nuat i on	( 40 o
c	23	330.2	43.7	110000	14	AE017225_39	Cont i nuat i on	( 40 o
c	24	330.2	43.7	110000	14	AE017334_39	Cont i nuat i on	( 40 o
c	25	330.2	43.7	110000	14	AE017355_38	Cont i nuat i on	( 39 o
c	26	330.2	43.7	110000	14	AE017355_39	Cont i nuat i on	( 40 o
c	27	329.4	43.6	1777	14	BACSP01 B	L47360	Bacillus st
c	28	328.6	43.5	759	8	EF039064	EF039064	Synt het ic
c	29	323.2	42.8	2073	14	BACSP02AB	M17643	B. subt illis
c	30	317	41.9	1907	14	BACSP01 A	M25260	Bacillus illi
c	31	317	41.9	110000	14	CP000002_24	Cont i nuat i on	( 25 o
c	32	317	41.9	110000	14	AE017333_24	Cont i nuat i on	( 25 o
c	33	316.2	41.8	110000	14	BA000028_18	Cont i nuat i on	( 19 o
c	34	273.4	36.2	665	2	AR805946	AR805946	Sequence
c	35	273.4	36.2	665	2	AX437345	AX437345	Sequence
c	36	254.6	33.7	1857	14	BACSP01 A	L47359	Bacillus sp
c	37	253.4	33.5	110000	14	CP000232_15	Cont i nuat i on	( 16 o
c	38	244.2	32.3	110000	14	CP000817_17	Cont i nuat i on	( 18 o
c	39	244.2	32.3	110000	14	CP000817_18	Cont i nuat i on	( 19 o
c	40	238.8	31.6	110000	14	AP009389_12	Cont i nuat i on	( 13 o
c	41	234.2	31.0	110000	14	CP000612_11	Cont i nuat i on	( 12 o
c	42	233.2	30.8	110000	14	CP000232_08	Cont i nuat i on	( 9 of
c	43	231	30.6	110000	14	CP000930_04	Cont i nuat i on	( 5 of
c	44	224.6	29.7	110000	14	AP008230_32	Cont i nuat i on	( 33 o
c	45	223.6	29.6	110000	14	CP000930_20	Cont i nuat i on	( 21 o

## ALIGMENTS

## RESULT 1

DD160143

LOCUS

DD160143 756 bp DNA linear PAT 23-NOV-2005  
DEFINITION Novel Brevibacillus choshinensis and Producing Method of Protein by  
Using Ther eof as Host .

ACCESSION

DD160143

VERSION

DD160143. 1 GI : 83960018

KEYWORDS

WO 2005045005- A/1.

SOURCE

Brevibacillus choshinensis

ORGANISM

Brevibacillus choshinensis

Bacteri a; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.

REFERENCE

1 ( bases 1 to 756)

AUTHORS

Hanagata, H. and Nishi jyo, T.

TITLE

Novel Brevibacillus choshinensis and Producing Method of Protein by  
Using Ther eof as Host

JOURNAL

Pat ent : WO 2005045005- A 1 19- MAY- 2005;

Higeta Shoyu Co Ltd

COMMENT

OS Brevibacillus choshinensis

PN WO 2005045005- A/1  
 PD 19- MAY- 2005  
 PF 08- NOV- 2004 WO 2004JP016912  
 PR 11- NOV- 2003 JP 03P 381606  
 PI hi r oshi hanagat a, t akayuki ni shi j yo  
 CC

FH Key Location/ Qualifiers  
 FEATURES Locat i on/ Qual i fiers  
 source 1. . 756  
 / organi sm=" Br evi baci ll us choshi nensi s"  
 / mol \_t ype=" unassi gned DNA"  
 / db\_xref =" t axon: 54911"

## ORIGIN

Query Match 100. 0% Score 756; DB 2; Length 756;  
 Best Local Similarity 100. 0% Pred. No. 6. 7e- 232;  
 Matches 756; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGGTCCGATATCAAAATGCGAGTCAACCATTCTGACCAATGACCAAGTGAAAGAT	60
Db	1	ATGGGTCCGATATCAAAATGCGAGTCAACCATTCTGACCAATGACCAAGTGAAAGAT	60
Qy	61	TTGATAGCCAAGAGCCAAGCTGGCGATAACGGATGCACGTGAGCTTCTCGTGAATAGCAAT	120
Db	61	TTGATAGCCAAGAGCCAAGCTGGCGATAACGGATGCACGTGAGCTTCTCGTGAATAGCAAT	120
Qy	121	ATCAGACTGGTCTGGTCCGTCGTCCAGCGCTTATCAACCGGGGTATGAAGCGGATGAT	180
Db	121	ATCAGACTGGTCTGGTCCGTCGTCCAGCGCTTATCAACCGGGGTATGAAGCGGATGAT	180
Qy	181	TTGTTTCAGATCGGTTGCATTGGCTTGTCAAGGCCGTTGACAAGTTGATCTTCTGTCAC	240
Db	181	TTGTTTCAGATCGGTTGCATTGGCTTGTCAAGGCCGTTGACAAGTTGATCTTCTGTCAC	240
Qy	241	GATGTGAGATTTGACCTATGCGGTGCGAATGATCATGGAGAAATTCAACGCTTTTG	300
Db	241	GATGTGAGATTTGACCTATGCGGTGCGAATGATCATGGAGAAATTCAACGCTTTTG	300
Qy	301	CGCGATGACGGTACGGTTAAGGTCACTCGATCGTTAAAAGAAACAGCGAATAAGGTGCGG	360
Db	301	CGCGATGACGGTACGGTTAAGGTCACTCGATCGTTAAAAGAAACAGCGAATAAGGTGCGG	360
Qy	361	CGATCAAAGGATGAATTGTACAAGCAATTGGCGGTGCGATCGCAGAAGTGGCA	420
Db	361	CGATCAAAGGATGAATTGTACAAGCAATTGGCGGTGCGATCGCAGAAGTGGCA	420
Qy	421	GAAGCAGTGGGAATCACGCGGGAGGAAGTAGTCGGCAAGAGGCAAGCAGAGCGCT	480
Db	421	GAAGCAGTGGGAATCACGCGGGAGGAAGTAGTCGGCAAGAGGCAAGCAGAGCGCT	480
Qy	481	TCTCCATCCATGAGACCGTTTTGAAAATGACGGCGATCCCATCACACTGATCGATCAG	540
Db	481	TCTCCATCCATGAGACCGTTTTGAAAATGACGGCGATCCCATCACACTGATCGATCAG	540
Qy	541	ATAGCGGATGAAGGTGTGAACAAGTGGTTGAGAAAATTGCGCTGAAGGAGGCCATCAGC	600
Db	541	ATAGCGGATGAAGGTGTGAACAAGTGGTTGAGAAAATTGCGCTGAAGGAGGCCATCAGC	600
Qy	601	AGGCTGAGCGAGCGTGAGCAGCTCATCGTCTACCTGCGCTATTACAAGGATCAGACACAG	660
Db	601	AGGCTGAGCGAGCGTGAGCAGCTCATCGTCTACCTGCGCTATTACAAGGATCAGACACAG	660
Qy	661	TCTGAGGTAGCAGAGCGCTAGGGATTGCGAGGTCCAGGCTCGCGTCTGGAAAAGCGT	720
Db	661	TCTGAGGTAGCAGAGCGCTAGGGATTGCGAGGTCCAGGCTCGCGTCTGGAAAAGCGT	720
Qy	721	ATCCTGCTAACGATCAAGGAGCAAATTGAACATTAG	756
Db	721	ATCCTGCTAACGATCAAGGAGCAAATTGAACATTAG	756

## RESULT 2

DI 115290

LOCUS DI 115290 756 bp DNA linear PAT 21- FEB- 2008  
 DEFINITION Novel *Brevibacillus choshinensis* and Producing Method of Protein by  
 Using Threonine as Host.

ACCESSION DI 115290

VERSION DI 115290. 1 GI : 168311716

KEYWORDS KR 1020067008845- A 1.

SOURCE *Brevibacillus choshinensis*ORGANISM *Brevibacillus choshinensis*REFERENCE Bacteria; Firmicutes; Bacillales; Paenibacillaceae; *Brevibacillus*.  
 1 (bases 1 to 756)

AUTHORS Hanagata, H. and Nishi jyo, T.

TITLE Novel *Brevibacillus choshinensis* and Producing Method of Protein by  
 Using Threonine as Host

JOURNAL Patent : KR 1020067008845- A 1 08- MAY- 2006;

COMMENT PN KR 1020067008845- A 1

PD 2006- 05- 08

PA HANAGATA, H., NISHI JYO, T.

TY DNA

OS *Brevibacillus choshinensis*

CO.

FEATURES Location/ Qualifiers

source 1..756

/organism="Brevibacillus choshinensis"

/mol\_type="unassigned DNA"

/db\_xref="taxon: 54911"

## ORIGIN

Query Match 100. 0% Score 756; DB 2; Length 756;  
 Best Local Similarity 100. 0% Pred. No. 6. 7e- 232;  
 Matches 756; Conservative 0; Missmatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGTGCCGATATCAAAATGCGAGTCAACCATTCTGACCAATGACCAAGTGAAAGAT 60  
 ||||| ||||| ||||| ||||| ||||| |||||

Db 1 ATGGGTGCCGATATCAAAATGCGAGTCAACCATTCTGACCAATGACCAAGTGAAAGAT 60  
 ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 TTGATAGCCAAGAGCCAAGCTGGCGATACGGATGCACGTGAGCTTCTCGTGAATAGCAAT 120  
 ||||| ||||| ||||| ||||| |||||

Db 61 TTGATAGCCAAGAGCCAAGCTGGCGATACGGATGCACGTGAGCTTCTCGTGAATAGCAAT 120  
 ||||| ||||| ||||| ||||| |||||

Qy 121 ATCAGACTGGTCTGGTCCGTCGTCAGCGCTTTATCAACCGCGGGTATGAAGCGGATGAT 180  
 ||||| ||||| ||||| ||||| |||||

Db 121 ATCAGACTGGTCTGGTCCGTCGTCAGCGCTTTATCAACCGCGGGTATGAAGCGGATGAT 180  
 ||||| ||||| ||||| ||||| |||||

Qy 181 TTGTTTCAGATCGGTTGCATTGGCTTGCTCAAGGCCGTTGACAAGTTGATCTTCGTAC 240  
 ||||| ||||| ||||| ||||| |||||

Db 181 TTGTTTCAGATCGGTTGCATTGGCTTGCTCAAGGCCGTTGACAAGTTGATCTTCGTAC 240  
 ||||| ||||| ||||| ||||| |||||

Qy 241 GATGTGAGATTTGACCTATGCGGTGCCAATGATCATCGGAGAAATTCAACGCTTTTG 300  
 ||||| ||||| ||||| ||||| |||||

Db 241 GATGTGAGATTTGACCTATGCGGTGCCAATGATCATCGGAGAAATTCAACGCTTTTG 300  
 ||||| ||||| ||||| ||||| |||||

Qy 301 CGCGATGACGGTAAGGTTAAGGTCACTCGATCGTTAAAAGAAACAGCGAATAAGGTGCGG 360  
 ||||| ||||| ||||| ||||| |||||

Db 301 CGCGATGACGGTAAGGTTAAGGTCACTCGATCGTTAAAAGAAACAGCGAATAAGGTGCGG 360  
 ||||| ||||| ||||| ||||| |||||

Qy 361 CGATCAAAGGATGAATTGACAAGCAATTGGCGGTGCCCCACGATCGCAGAAGTGGCA 420  
 ||||| ||||| ||||| ||||| |||||

Db 361 CGATCAAAGGATGAATTGACAAGCAATTGGCGGTGCCCCACGATCGCAGAAGTGGCA 420  
 ||||| ||||| ||||| ||||| |||||

Qy 421 GAAGCACTGGGATCACGCCGGAGGAAGTAGTCTTGCGCAAGAGGCAACGAGGCGCT 480  
 ||||| ||||| ||||| ||||| |||||

Db 421 GAAGCACTGGGATCACGCCGGAGGAAGTAGTCTTGCGCAAGAGGCAACGAGGCGCT 480  
 ||||| ||||| ||||| ||||| |||||

Qy 481 TCCTCCATCCATGAGACCGTTTGAAAATGACGGCGATCCCATCACACTGATCGATCAG 540  
 ||||| ||||| ||||| ||||| |||||

Db	481	TCCTCCATCCATGAGACCGTTTTGAAAATGACGGCGATCCCATCACACTGATCGATCAG	540
Qy	541	ATAGCGATGAAGGTGTGAACAAGTGGTTGAGAAAATTGCCTGAAGGACGCCATCAGC	600
Db	541	ATAGCGATGAAGGTGTGAACAAGTGGTTGAGAAAATTGCCTGAAGGACGCCATCAGC	600
Qy	601	AGGCTGAGCGAGCGTGAGCAGCTCATCGTCTACCTGCGCTATTACAAGGATCAGACACAG	660
Db	601	AGGCTGAGCGAGCGTGAGCAGCTCATCGTCTACCTGCGCTATTACAAGGATCAGACACAG	660
Qy	661	TCTGAGGTAGCAGAGCGCTAGGGATTTCGCAGGTCCAGGTCTCGCGTCTGGAAAAGCGT	720
Db	661	TCTGAGGTAGCAGAGCGCTAGGGATTTCGCAGGTCCAGGTCTCGCGTCTGGAAAAGCGT	720
Qy	721	ATCCTGCTAACGATCAAGGAGCAAATTGAACATTAG	756
Db	721	ATCCTGCTAACGATCAAGGAGCAAATTGAACATTAG	756

## RESULT 3

BACSP01

LOCUS BACSP01 1746 bp DNA linear BCT 16-OCT-1997

DEFINITION Paenibacillus polymyxa spolIAA, spolIAB, and sigma factor genes, complete cds.

ACCESSION L47358

VERSION L47358.1 G : 2529264

KEYWORDS

SOURCE Paenibacillus polymyxa

ORGANISM Paenibacillus polymyxa

REFERENCE Bacteriophages; Bacteriophages; Paenibacillaceae; Paenibacillus.

AUTHORS Park, S. G. and Yudkin, M. D.

TITLE Sequencing and phylogenetic analysis of the spolI A operon from diverse *Bacillus* and *Paenibacillus* species

JOURNAL Gene 194 (1), 25-33 (1997)

PUBMED 9266669

COMMENT GSDB: S: 46599.

FEATURES

source Location/Qualifiers

1..1746 /organism="Paenibacillus polymyxa" /mol\_type="genomic DNA" /db\_xref="taxon:1406" /clone\_id="DSM86"

-35\_signal 6..11 /inference="non-experimental evidence, no additional details recorded"

-10\_signal 29..33 /inference="non-experimental evidence, no additional details recorded"

gene 70..423 /gene="spolI AA"

CDS 70..423 /gene="spolI AA" /inference="non-experimental evidence, no additional details recorded" /codon\_start=1 /transl\_table=11 /protein\_id="AAB81184.1" /db\_xref="GI:2529265" /translation="MNLQI EMEHHRGVLI VRLSGELDHHTSDMVRMQMDEAI QRRQCE HI VL SLKNLQFMDSSGLGV LGYKLI NQKGGEAVCDVNPPVHRLLDMSGLFKI MPI YDNEVNALTELEVVS"

gene 420..875 /gene="spolI AB"

CDS 420..875 /gene="spolI AB" /inference="non-experimental evidence, no additional details recorded" /codon\_start=1

/transl\_table=11  
 /protein\_id="AAB81185.1"  
 /db\_xref="GI:2529266"  
 /translation="MREGTGNFMSLQFAAKSENAAFARVAVAIFI SRLDPTMDELSD  
 LKTVVSEAVNSI I HGYDSDPSGVVTI KVGIEADVI TLVVEDAGRGI EDLELAQPLY  
 TSKPELERSGMGFTI MENFMDFEAVSEPGGGTSVRMKKRI ESKKALYN"  
 CDS 887..1642  
 /inference="non-experimental evidence, no additional  
 detailed score"  
 /codon\_start=1  
 /transl\_table=11  
 /product="sigma factor"  
 /protein\_id="AAB81186.1"  
 /db\_xref="GI:2529267"  
 /translation="MEAGGKKTSHSYLEDTEVKRLI ALSQSGDNDARETLI NSNI RLV  
 WSVWQRFMNRGYEPDDL FQI GCI GLLKSVDKFDLSYEVKFSTYAVPM I GEI QRFLRD  
 DGTLKVSRLSLKETANKVRKMKDELSKPLNRLPTVKEVADEL GVTPEDWVFAQEANKPP  
 TSI HETVFENDGDPITLMDQI ADESQERWFDFKLALNEAI DGLSERERLI VYLRYYRDQ  
 TQSEVASRLGI SQVQVSRLLEKKI LQQI RDQI AQ"

## ORIGIN

Query Match 48.4% Score 366.2; DB 14; Length 1746;  
 Best Local Similarity 67.8% Pred. No. 5.2e-106;  
 Matches 512; Conservative 0; Missmatches 243; Indels 0; Gaps 0;

Qy	1	ATGGGTGCCGATATCAAAATGCGAGTCACCATTTCTGACCAATGACCAAGTGAAAGAT	60
Db	887	ATGGAACCAGGAGGAAAAAAACTTCGCACAGCTATTGGAGGATACAGAAGTCAGCGG	946
Qy	61	TTGATAGCCAAGAGCCAAGCTGGCGATACGGATGCCACGTGAGCTTCTCGTGAATAGCAAT	120
Db	947	CTTATGCCACTGAGTCATCTGGTATAATGATGCTCGTGAGACACTAATCAATAGTAAT	1006
Qy	121	ATCAGACTGGCTGGTCCGTCGTCAGCGCTTTATCAACCGGGGTATGAAGCGGATGAT	180
Db	1007	ATCCGGCTCGTCTGGTCTGTCGTCAGCGCTTTATGAACCGGGGTATGAACCTGACGAT	1066
Qy	181	TTGTTTCAGATCGGTTGCATTGGCTTGCTCAAGGCGGTTGACAAGTTGATCTTCTGAC	240
Db	1067	CTTTTCAAATTGGTGCATCGGATTGTTAAATCGTGGATAATTGATCTCAGCTAC	1126
Qy	241	GATGTGAGATTTCGACCTATGGGTGCCAATGATCATGGAGAAATTCAACGCTTTTG	300
Db	1127	GAAGTGAAGTTCTCATATGCAAGTACCAATGATTATTGGTGGATTCAACGCTCCCT	1186
Qy	301	CGCGATGACGGTACGGTAAGGTCACTCGATCGTTAAAAGAAACAGCGAATAAGGTGGG	360
Db	1187	CGAGACGATGGTACTCTCAAGTTAGCGTTCCCTGAAGGAAACAGCGAATAAGGCCGC	1246
Qy	361	CGATCAAAGGATGAATTGACAAGCAATTGGCGGTGCCCCACGATCGCAGAAGTGGCA	420
Db	1247	AAAATGAAGGACGAGCTGTCAAACCGCTCAATCGCTGCCTACAGTCAGGAGTTGCT	1306
Qy	421	GAAGCAGTGGGATCACGCGGGAGGAAGTAGTCCTTGCGCAAGAGGCAAGCAGAGGCCT	480
Db	1307	GATGAGCTGGAGTGCACACGGAGGATGTTGTTTGCGCAGGAGGCCATAAGCCACCT	1366
Qy	481	TCCTCCATCCATGAGACCGTTTGTAAAATGACGGCGATCCCATCACACTGATCGATCAG	540
Db	1367	ACCTCCATTACGAAACAGTGTGTTGAAAATGACGGAGACCGATTACACTTATGGATCAG	1426
Qy	541	ATAGCGGATGAAGGTGTGAAACAAGTGGTTGAGAAAATTGCGTTGAAGGACGOCATCAGC	600
Db	1427	ATTGCCGACGAATCCAGGAAAGATGGTTCGACAAGCTGGCGTGAATGAGGCCATTGAC	1486
Qy	601	AGGCTGAGCGAGCGTGAGCAGCTCATCGTCTACCTGCGCTATTACAAGGATCAGACACAG	660
Db	1487	GGATTATCTGAACGGGAGCGATTGATTGTGTAACCTGCGTTATTACCGGGATCAGACTCAG	1546

Qy 661 TCTGAGGTAGCAGAGCGCTAGGGATTTCGCAGGTCCAGGCTCGCGTCTGGAAAAGCGT 720  
 Db 1547 TCAGAGGTTGCAAGTCGCTTGGCATTCTCAGGTCCAGGTATCAGGATTGGAAAAAAAG 1606  
 Qy 721 ATCCTGCTAACGATCAAGGAGCAAATTGAACATTA 755  
 Db 1607 ATTTTACAACAGATCCGCGATCAGATTGCCAATA 1641

## RESULT 4

Z54161  
 LOCUS Z54161 1799 bp DNA linear BCT 18- APR- 2005  
 DEFINITION B. coagulans genes for Spol I AA, Spol I AB and sigma factor.  
 ACCESSION Z54161  
 VERSION Z54161. 1 GI : 1552541  
 KEYWORDS sigma factor; Spol I AA; Spol I AB.  
 SOURCE Bacillus coagulans  
 ORGANISM Bacillus coagulans  
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 REFERENCE 1 (bases 1 to 1799)  
 AUTHORS Park, S. G. and Yudkin, M D.  
 TITLE Nucleotide sequence of the *Bacillus coagulans* homologue of the spol I A operon of *Bacillus subtilis*  
 JOURNAL Gene 177 (1-2), 275-276 (1996)  
 PUBMED 8921882  
 REFERENCE 2 (bases 1 to 1799)  
 AUTHORS Park, S. S.  
 TITLE Direct Submission  
 JOURNAL Submitted (08- SEP- 1995) Sung Goo SG Park, Microbiology Unit,  
 Biochemistry, University of Oxford, South Parks Road, Oxford, OX1  
 3QU, UK  
 FEATURES source  
 LOCATION/QUALIFIERS  
 1..1799  
 /organism="Bacillus coagulans"  
 /mol\_type="genomic DNA"  
 /isolate="DSM1"  
 /db\_xref="taxon:1398"  
 -35\_signal 75..80  
 -10\_signal 98..101  
 CDS 137..487  
 /not e="gt g start codon"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Spol I AA"  
 /protein\_id="CAA90872.1"  
 /db\_xref="GI:1552542"  
 /db\_xref="GOA: P70877"  
 /db\_xref="InterPro: IPR002645"  
 /db\_xref="InterPro: IPR003658"  
 /db\_xref="UniProtKB/Swiss-Prot: P70877"  
 /translation="MSLAI DLEVKKDVLCI RLQGELDHHTAESLRASVTKAI EENGL R  
 HLVLNL EQLSFMDSSGLGVI LGRYKQI KQKNGEM VCAI SPAVKRLFELSGLFKI I RL  
 DESEQYALHRLGVA"  
 CDS 488..931  
 /codon\_start=1  
 /transl\_table=11  
 /product="Spol I AB"  
 /protein\_id="CAA90873.1"  
 /db\_xref="GI:1552543"  
 /db\_xref="GOA: P70878"  
 /db\_xref="InterPro: IPR003594"  
 /db\_xref="InterPro: IPR010194"  
 /db\_xref="UniProtKB/Swiss-Prot: P70878"  
 /translation="MKNKMQLQFTARSNESFARVRVAAFVAQLDPTLDELTGQNTVV  
 SEAVTNALI HGYEENPDGMVYVSAAI CDDGTVEI TVRDEGKGADI EEARQPSFTTKP  
 ELERAGMGFTI MENFMDDEI QSQQESGTI I RLKKHLALKNALCN"  
 CDS 944..1702  
 /codon\_start=1  
 /transl\_table=11

```

/ pr oduct ="si gma_ f act or "
/ pr ot ei n_i d="CAA90874. 1"
/ db_xr ef ="GI : 1552544"
/ db_xr ef ="GOA: P70879"
/ db_xr ef ="HSSP: P00579"
/ db_xr ef ="I nt er Pr o: I PR000943"
/ db_xr ef ="I nt er Pr o: I PR007624"
/ db_xr ef ="I nt er Pr o: I PR007627"
/ db_xr ef ="I nt er Pr o: I PR007630"
/ db_xr ef ="I nt er Pr o: I PR013324"
/ db_xr ef ="I nt er Pr o: I PR013325"
/ db_xr ef ="Uni Pr ot KB/ Tr EMBL: P70879"
/ t ransl at i on="MDAEL KKAENKARL KDEEVKNLI RRTEEGDQEARDL I VEKNI RL
VMSVVRFLNRYEPDDL FQI GCI GLLKSVDKFDSL SYDVKFSTYAVPM I GEI QRFLR
DDGTVKVSRSLKETGNRI RKI RDEL TKKLGRSPTVHEI ASELGMPVEEVVMAQEAGKA
PSSI HETVYENDGDPI TLLDQI SDQNDQKWF DQI ALQEAI RGLEERERLI VYLRYYKD
QTQSEAGERLGI SQVQVSRL EKKI LRVMKERML"

```

## ORI G N

Query	Match	48. 0%	Score	363. 2;	DB	14;	Length	1799;	
Best Local Similarity		69. 4%	Pr ed. No.	4. 9e- 105;					
Matches	494;	Conservative	0;	M smatches	218;	Indels	0;	Gaps	0;
Qy	44	ATGACCAAGT GAAAGATT GATAGCCAAGAGCCAAGCTGGCGATA CGGATGCACGTGAGC	103						
Db	990	ATGAGGAAGT GAAAAACCT GATCAGGGCACCGAGGAAGGCGACCAGGAGGCACGCGACC	1049						
Qy	104	TTCTCGTGAATAGCAATATCAGACTGGTCTGGTCCGT CGTCCAGCGCTTTATCAACCGCG	163						
Db	1050	TGATCGTTGAAAAAAATATCCGCCTTGTCTGGTCCGGTTGTCCACCGCTTTAAACCGGG	1109						
Qy	164	GGTATGAAGCGGATGATTGTTT CAGATCGGTTGCATTGGCTTGCTCAAGGCGTTGACA	223						
Db	1110	GATATGAACCGGATGACTTATTCAAATCGGCTGTATCGGCTGTAAAATCGGTGATA	1169						
Qy	224	AGTTCGATCTT CGTACGATGTGAGATTT CGACCTATCGGGTGCCAATGATCATCGGAG	283						
Db	1170	AATTGATCTT CGTATGATGTCAAATTTCACCTATGCAGTCCCCATGATCATCGGGG	1229						
Qy	284	AAATTCAACGCTTTT GCGCGATGACGGTACGGTAAGGT CAGTCATCGTTAAAGAAA	343						
Db	1230	AAATCCAGCGGTTTT GCGCGATGACGGGACGGTAAGGT GAGCGTTCTTGAAAGAAA	1289						
Qy	344	CAGCGAATAAGGT GCGCGATCAAAGGATGAATTGTACAAGCAATTGGCGGTGCCCCCA	403						
Db	1290	CGGGCAACCGATCCGCAAATCCCGATGAGTTAACGAAAAGCTGGGAGATCGCCGA	1349						
Qy	404	CGATCGCAGAAGT GGCAGAACGAGTGGAAATCACGCCGGAGGAAGTAGTCCTTGCGCAAG	463						
Db	1350	CCGTGCATGAAATTGCATCAGAACTCGGATGCCGGTGGAGAAGTCGTATGGCGCAGG	1409						
Qy	464	AGGCAACCAGAGCGCTT CCTCCATCCATGAGACCGTTTTGAAAATGACGGCGATCCC	523						
Db	1410	AGGCAGGCAAAGCCCCATCATCTATCCATGAGACGGTATATGAAAACGATGGCGACCGA	1469						
Qy	524	TCACACTGATCGATCAGATAGCGGTGAAGGTGTGAACAAGTGGTTGAGAAAATTGCC	583						
Db	1470	TCACGCTCTTGACCAGATT CAGACCGAGACGACGAGAAGTGGTTGACCAGATTGCC	1529						
Qy	584	TGAAGGACGCCATCAGCAGGCTGAGCGAGCGT GAGCAGCTCATCGTCACTCGCTATT	643						
Db	1530	TCCAGGAAGCGATCAGAGGGCTTGAGGAACGCGAAAGGCTGATTGTGACCTCGCTATT	1589						
Qy	644	ACAAGGATCAGACACAGTCTGAGGTACCGAGCGCTAAGGATTTCCAGGTCCAGGTCT	703						
Db	1590	ACAAAGACCAACACAGTCGAAGCTGGGAACGGCTGGCATT CGCAGGTCCAGGTCT	1649						
Qy	704	CGCGTCTGGAAAAGCGTATCCTGCTAACGATCAAGGAGCAAATTGAACATTA	755						

Db 1650 CACGGCTT GAAAAGAAAATCTCGAGTGATGAAAGAGOGCATGGAATTGTA 1701

RESULT 5  
 AJ704768  
 LOCUS AJ704768 3701 bp DNA linear BCT 10-MAY-2005  
 DEFINITION Bacillus amylolyquefaciens dacF gene, spoI<sub>A</sub> operon and spoVAA gene (partial), strain FZB42.  
 ACCESSION AJ704768  
 VERSION AJ704768.1 GI:63146678  
 KEYWORDS anti-sigma F factor; anti-sigmaF factor; dacF gene; penicillin binding protein; sigF gene; Sigma F; spoI<sub>AA</sub> gene; spoI<sub>AB</sub> gene; spoVAA gene; Stage V sporulation protein AA.  
 SOURCE Bacillus amylolyquefaciens FZB42  
 ORGANISM Bacillus amylolyquefaciens FZB42  
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 REFERENCE 1  
 AUTHORS Strohinger, P.  
 TITLE Comparison of early sporulation genes in Bacilli and Clostridia  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 3701)  
 AUTHORS Borriiss, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (10-MAY-2004) Borriiss R., Biology, Humboldt University, Chaussee Strasse 117, D-10115, GERMANY  
 COMMENT spoI<sub>A</sub> operon 1290 - 3083.  
 FEATURES source  
 gene  
 CDS  
 operon  
 gene  
 CDS

Location/Qualifiers  
 1..3701  
 /organism="Bacillus amylolyquefaciens FZB42"  
 /mol\_type="genomic DNA"  
 /strain="FZB42"  
 /isolation\_source="soil"  
 /db\_xref="Taxon: 326423"  
 /country="Germany: Berlin-Brandenburg"  
 123..1292  
 /gene="dacF"  
 123..1292  
 /gene="dacF"  
 /EC\_number="3.4.16.4"  
 /function="required for spore cortex synthesis"  
 /note="putative D-Ala-D-Ala carboxypeptidase"  
 /codon\_start=1  
 /transl\_table=11  
 /product="penicillin binding protein"  
 /protein\_id="CAG28930.1"  
 /db\_xref="GI:63146679"  
 /db\_xref="GOA:Q50J84"  
 /db\_xref="HSSP:P39042"  
 /db\_xref="InterPro: IPR001967"  
 /db\_xref="InterPro: IPR012338"  
 /db\_xref="InterPro: IPR012907"  
 /db\_xref="UniprotKB/TrEMBL: Q50J84"  
 /translation="MKRLISTLLIGILLTASAPSAFKPDGKHTSELAEAKSAVLI ERDTGSLYNKNSRERLAPASMTKIMTMLI MEALDKGKI KMSDKVRTSEHAASMGGSQ I FLEPGEEMTVKEMLKGI AI ASGNDASVAMAEGI AGSEEDFVSRMNKKAKELGLKDT S FKNPTGLPEKDHSSAYDMAKMAKELLKYDKI TGFTGTYEDYLRENTDKKFWLVNNTNR LI KFYPGVDGVTGFTGEAKYCLTATAKKGMRVI AVVFGASTPKERNAQVTKMLDYA FSQFKTHPLYKRDQI VGTVKVKGKQKLI KLTTSEPI SLLAKKGENMDVKKEVKI NG NVTAPVKGEVLGSLVLKKDGVLVESPVTAKDDMEKAGFLTFLKRTMDWTKFK"  
 1290..3083  
 /operator="spoI<sub>A</sub>"  
 /function="expression of early forespore genes"  
 1388..1741  
 /gene="spoI<sub>AA</sub>"  
 /operator="spoI<sub>A</sub>"  
 1388..1741  
 /gene="spoI<sub>AA</sub>"  
 /operator="spoI<sub>A</sub>"  
 /function="binding to SpoI<sub>AB</sub> in the presence of ADP"

```

selectively in the forespore"
/codon_start=1
/translate_table=11
/product="anti-sigma F factor antagonist"
/protein_id="CAG28931.1"
/db_xref="GI:63146680"
/db_xref="HSSP: P10727"
/db_xref="InterPro: IPR002645"
/db_xref="InterPro: IPR003658"
/db_xref="UniProtKB/TrEMBL: Q50J83"
/translation="MSLGI DFHVKESVLCI RLTGELDHITAESL KQKVTESELADGI R
HI VL NLEDLT FMDSGGLGVI LGRYKQI KQLGGMVCAI SPAVRRLF DMSGLFKI I RF
EQSEQKALLTLCVVAS"
gene 1738. . 2178
/gene="spolIAB"
/operon="spolIAB"
CDS 1738. . 2178
/gene="spolIAB"
/operon="spolIAB"
/EC_number="2.7.1.37"
/function="binding to sigmaF in the presence of ATP before
separation"
/codon_start=1
/translate_table=11
/product="anti-sigmaF factor"
/protein_id="CAG28932.1"
/db_xref="GI:63146681"
/db_xref="GOA: Q50J82"
/db_xref="HSSP: C32727"
/db_xref="InterPro: IPR003594"
/db_xref="InterPro: IPR010194"
/db_xref="UniProtKB/TrEMBL: Q50J82"
/translation="MKNEMHLEFSALSQNESFARVTVASFI AQLDPTMDELTEI KTVW
SEAVTNALI HGYEENCDGKVYI SVTLEDHVYVLTIRDEGI ITDLEEARQPLFTTKPE
LERSGMGFTI MENFMDVTI DSSPEMGTI RLTKHLSSKALCN"
gene 2190. . 2957
/gene="sigF"
/operon="spolIAB"
CDS 2190. . 2957
/gene="sigF"
/operon="spolIAB"
/function="sporulation specific sigma factor"
/codon_start=1
/translate_table=11
/product="Sigma F"
/protein_id="CAG28933.1"
/db_xref="GI:63146682"
/db_xref="GOA: Q50J81"
/db_xref="HSSP: P00579"
/db_xref="InterPro: IPR000943"
/db_xref="InterPro: IPR007624"
/db_xref="InterPro: IPR007627"
/db_xref="InterPro: IPR007630"
/db_xref="InterPro: IPR013325"
/db_xref="UniProtKB/TrEMBL: Q50J81"
/translation="MDVEVKKN SKNAQ LKDHEVKE LKKSQEGDQQARDL LI EKNMRL
VWSVVQRFL NRGYEPDDL FQI GCI GLLKSVDKF DLSYDVRF STYAVPM I GEI QRFI R
DDGTVKVSRSL KELGNKI RRAKDEL SKTYGRVPTVQE I ADHLEI DAEDVWLAQEAVRA
PSSI HETVYENDGDP I TLLDQI ADHSEEKWF DKI ALKEAI SDLEEREKLI VYLRYYKD
QTQSEVAERLGI SQVQVSRL EKKI LKQI KVQMDHTES"
gene 3084. . >3701
/gene="spoVAA"
CDS 3084. . >3701
/gene="spoVAA"
/codon_start=1
/translate_table=11
/product="Stage V sporulation protein AA"
/protein_id="CAG28934.1"
/db_xref="GI:63146683"

```

/ db\_xr\_ef = "UniProtKB/TrEMBL: Q50J80"  
 / translation="MERRIFLRLRHRVKAHTGDIITI GDAAQIAGQQKLKKKLSALPL  
 YKVTKDKNIVVLDIIVKVLKTIHTAFFALDVQTVGGAETIVEIHYQKKQLSAVLFVGV  
 VLLLFI GSCLAI MNFHEDVSMREVHIALYEIITGERNEYPYLLQI PYSI GLGLGMVF  
 FNHI FKKRLNEEPSLEVENMFNYQLLDHYVAMHENEETMKDLHDR"

## ORIGIN

Query Match 47.8% Score 361.6; DB 14; Length 3701;  
 Best Local Similarity 67.9% Pred. No. 1.8e-104;  
 Matches 505; Conservative 0; Mismatches 239; Indels 0; Gaps 0;

Qy	10	GATATCAAAAATGCGAGTCACCAATTCTGACCAATGACCAAGTGAAAGATTGATGCC	69
Db	2202	GTAAAGAAAAACAGCAAAACGCGCAGCTCAAGGACCATGAAGTCAAAGAATTAACTAAA	2261
Qy	70	AAGAGCCAAGCTGGCGATACGGATGCAACGTGAGCTTCTCGTGAATAGCAATATCAGACTG	129
Db	2262	AAGAGCCAAGAGGGCGACCAGCAGCCAAAGAGACCTCCTCATAGAAAAAAACATGCGTCTT	2321
Qy	130	GTCTGGTCGGTCCAGGCTTTATCAACCGGGGTATGAAGGGATGATTGTTTCAG	189
Db	2322	GTGTGGTCGGTCCGGTCTTAAACAGAGGCTATGAGCCGGATGATCTGTTTCAG	2381
Qy	190	ATCGGTTGCATTGGCTTGCTCAAGCCCGTTGACAAGTTGATCTTCGATCGATGTGAGA	249
Db	2382	ATCGGCTGTATCGGATTGTTAAACATCGGTTGATAATTGATTTGTCCTATGATGCCGT	2441
Qy	250	TTTCGACCTATGCGGTGCGAATGATCATCGGAGAAATTCAACGCTTTGCCGATGAC	309
Db	2442	TTTCAACTTATGCCGTGCGATGATTATCGGAGAAATCCAGCCCTCATCCCCGATGAC	2501
Qy	310	GGTACGGTTAAGGTAGTCGATCGTTAAAGAAACAGCGAATAAGGTGGCGATCAAAG	369
Db	2502	GGAACCGTCAAAGTTCAAGATCTTAAAGAGCTGGCAATAAAATACGCCGGGGAAA	2561
Qy	370	GATGAATTGTACAAGCAATTGGCCGTGCCCCCACGATCGAGAAGTGGCAGAAGCAGTG	429
Db	2562	GATGAACGTCAAACATACGGAGAGTGGCGACCGTCCAGGAAATCGGGACCATTTA	2621
Qy	430	GGAATCACGCCGGAGGAAGTAGTCGGCAAGAGGCAAGCAGAGGCCCTCCATC	489
Db	2622	GAAATTGATGCCGGAGGATGTGGTCTGGCAGGAAGCGGTGCCGGGCCCTCATCCATT	2681
Qy	490	CATGAGACCGTTTGAAATGACGGCGATCCCATCACACTGATCGATCAGATAGCGGAT	549
Db	2682	CATGAGACCGTGTATGAAATGACGGAGACCGGATCACCTGCTGATCAGATGCCGAT	2741
Qy	550	GAAGGTGTGAACAAGTGGTTGAGAAAATTGCCCTGAAGGACGCCATCAGCAGGCTGAGC	609
Db	2742	CATTAGAGAGAAAGTGGTTGACAAAATGCCCTCAAGGAGGCGATCAGTGATTGGAA	2801
Qy	610	GAGCGTGAGCAGCTACGTCACCTGCGCTATTACAAGGATCAGACACAGTCAGGTA	669
Db	2802	GAACGGGAAACTCATTGCTATCTGAGATTAAAGATCAGACACAGTCGGAGTA	2861
Qy	670	GCAGAGCGTCTAGGGATTTCGCGAGGCTCAGGTCTCGCGTCTGGAAAAGCGTATCCTGCTA	729
Db	2862	GCGGAACGGCTCGGGATTCCAGGTCCAGGTATCGAGGCTGGAGAAAAAAATTAAAG	2921
Qy	730	ACGATCAAGGAGAAATTGAACAT	753
Db	2922	CAGATTAAAGTGCAGATGGATCAT	2945

## RESULT 6

CP000560\_22/c

WPCOMMENT

Sequence split into 40 fragments LOCUS CP000560 Accession CP000560  
 Fragment Name Begin End

CP000560_00	1	110000
CP000560_01	100001	210000
CP000560_02	200001	310000
CP000560_03	300001	410000
CP000560_04	400001	510000
CP000560_05	500001	610000
CP000560_06	600001	710000
CP000560_07	700001	810000
CP000560_08	800001	910000
CP000560_09	900001	1010000
CP000560_10	1000001	1110000
CP000560_11	1100001	1210000
CP000560_12	1200001	1310000
CP000560_13	1300001	1410000
CP000560_14	1400001	1510000
CP000560_15	1500001	1610000
CP000560_16	1600001	1710000
CP000560_17	1700001	1810000
CP000560_18	1800001	1910000
CP000560_19	1900001	2010000
CP000560_20	2000001	2110000
CP000560_21	2100001	2210000
CP000560_22	2200001	2310000
CP000560_23	2300001	2410000
CP000560_24	2400001	2510000
CP000560_25	2500001	2610000
CP000560_26	2600001	2710000
CP000560_27	2700001	2810000
CP000560_28	2800001	2910000
CP000560_29	2900001	3010000
CP000560_30	3000001	3110000
CP000560_31	3100001	3210000
CP000560_32	3200001	3310000
CP000560_33	3300001	3410000
CP000560_34	3400001	3510000
CP000560_35	3500001	3610000
CP000560_36	3600001	3710000
CP000560_37	3700001	3810000
CP000560_38	3800001	3910000
CP000560_39	3900001	3918589

Continuation (23 of 40) of CP000560 from base 2200001 (CP000560 *Bacillus amyloliquefaciens* AC

Query Match 47.8% Score 361.6; DB 14; Length 110000;  
 Best Local Similarity 67.9% Pred. No. 3.1e-104;  
 Matches 505; Conservative 0; Missmatches 239; Indels 0; Gaps 0;

Qy	10	GATATCAAAATGCGAGTCACCAATTCTGACCAATGACCAAGTGAAAGATTGATAGCC	69
Db	47250	GTAAAGAAAAACAGCAAAACGCGCAGCTAAGGACCATGAAGTCAAAGAATTAAATCAA	47191
Qy	70	AAGAGCCAAGCTGGGATAACGGATGAACTGAGCTTCTCGTGAATAGCAATATCAGACTG	129
Db	47190	AAGAGCCAAGAGGGGACCAGCAGCCAGCAAGAGACCTCCTCATAGAAAAACATGGCTCTT	47131
Qy	130	GTCGGTCCGTCGTCCAGCGCTTTATCAACCGCCGGTATGAGCGGATGATTGTTCAAG	189
Db	47130	GTCGGTCCGTCGTCAAAGGTTCTAAACAGAGGCTATGAGCGGATGATCTGTTCAAG	47071
Qy	190	ATCGGTTGCATTGGCTTGCTCAAGGCCGTTGACAAGTTGATCTTGTACGATGTGAGA	249
Db	47070	ATCGGCTGTATCGGATTGTTAAATCGGTTGATAAAATTGATTTGTCCTATGATGTCGGT	47011
Qy	250	TTTCGACCTATGCGGTGCGAATGATCATCGGAGAAATTCAACGCTTTGGCGATGAC	309
Db	47010	TTTCAACTATGCGGTGCGAATGATCATCGGAGAAATTCAACGCTTTGGCGATGAC	46951
Qy	310	GGTACGGTTAAGGTCAGTCGATCGTTAAAGAAACAGCGAATAAGGTGCGGCGATCAAAG	369
Db	46950	GGAACCGTCAAAGTTCAAGATCTTAAAGAGCTGGCAATAAAATACGCGGGCGAA	46891

Qy	370	GATGAATTGTACAAGCAATTGGCGGTGCCCCACGGATGCAGAAGTGGCAGAAGCAGTG	429
Db	46890	GATGAACGTCAAAACATACGGAAAGAGTGCAGACCGTCCAGGAAATCGGGACCATTTA	46831
Qy	430	GGAATCACGCCGGAGGAAGTAGTCCTTGCAGAAGAGGCAAGCAGAGGCCCTCCATC	489
Db	46830	GAAATTGATGCCGAGGATGTGGTGCCTGCAGGAAGGGTGCAGGCGCTTCATCCATT	46771
Qy	490	CATGAGACCGTTTGAAATGACGGCGATCCATCACACTGATCGATCAGATAGGGAT	549
Db	46770	CATGAGACCGTGTATGAAATGACGGAGACCGATCACCCGCTTGCATCAGATCGCGAT	46711
Qy	550	GAAGGTGTAAACAAGTGGTTGAGAAAATTGCCTTGAAGGACGOCATCAGCAGGCTGAGC	609
Db	46710	CATTAGAAAGAGAAAGTGGTTGACAAAATGCCCTCAAGGAGGGATCAGTGTGTTGGAA	46651
Qy	610	GAGCGTGGCAGCTCATCGTCTACCTGCGCTATTACAAGGATCAGACACAGTCTGAGGTA	669
Db	46650	GAACGGAAAAACTCATTGTCTATCTGAGATATTAAAGATCAGACACAGTOGGAAGTA	46591
Qy	670	GCAGAGCGTCTAGGGATTTCGCAGGTCCAGGTCTCGCGTCTGGAAAAGCGTATCCTGCTA	729
Db	46590	GCCGAACGGCTCGGATTCCCAAGTCCAGGTATGAGGCTGGAGAAAAAAATTAAAG	46531
Qy	730	ACGATCAAGGAGCAAATTGAACAT	753
Db	46530	CAGATTAAAGTGCAGATGGATCAT	46507

## RESULT 7

CP000813\_21/c

WPCOMMENT

Sequence split into 37 fragments LOCUS CP000813 Accession CP000813

Fragment	Name	Begin	End
CP000813_00		1	110000
CP000813_01		100001	210000
CP000813_02		200001	310000
CP000813_03		300001	410000
CP000813_04		400001	510000
CP000813_05		500001	610000
CP000813_06		600001	710000
CP000813_07		700001	810000
CP000813_08		800001	910000
CP000813_09		900001	1010000
CP000813_10		1000001	1110000
CP000813_11		1100001	1210000
CP000813_12		1200001	1310000
CP000813_13		1300001	1410000
CP000813_14		1400001	1510000
CP000813_15		1500001	1610000
CP000813_16		1600001	1710000
CP000813_17		1700001	1810000
CP000813_18		1800001	1910000
CP000813_19		1900001	2010000
CP000813_20		2000001	2110000
CP000813_21		2100001	2210000
CP000813_22		2200001	2310000
CP000813_23		2300001	2410000
CP000813_24		2400001	2510000
CP000813_25		2500001	2610000
CP000813_26		2600001	2710000
CP000813_27		2700001	2810000
CP000813_28		2800001	2910000
CP000813_29		2900001	3010000
CP000813_30		3000001	3110000
CP000813_31		3100001	3210000
CP000813_32		3200001	3310000
CP000813_33		3300001	3410000

CP000813\_34 3400001 3510000  
 CP000813\_35 3500001 3610000  
 CP000813\_36 3600001 3704465

Continuation (22 of 37) of CP000813 from base 2100001 (CP000813 *Bacillus pumilus* SAFR-

Query Match 46.9% Score 354.6; DB 14; Length 110000;  
 Best Local Similarity 67.6% Pred. No. 5.6e-102;  
 Matches 498; Conservative 0; Missmatches 239; Indels 0; Gaps 0;

Qy 16 AAAATGCGAGTCACCATTCTGACCAATGACCAAGTGAAAGATTGATAGCCAAGAGC 75  
 Db 14501 AAGCAACCAAAAAAGGCCAGCTGTCAAATGATGAAGTCAAAGAGCTCATCAAAAAAGC 14442

Qy 76 CAAGCTGGGATACGGATGCACTGAGCTCTCGTGAATAGCAATATCAGACTGGCTGG 135  
 Db 14441 CAAGACGGTGATCAGCAAGCAAGAGACCTCCTCGTAGAAAAAAACATGGCTTGTGG 14382

Qy 136 TCCGTCGTCCAGCGCTTATCAACCGGGGTATGAAGGGATGATTGTTTCAAGATCGGT 195  
 Db 14381 TCCGTTGTTCAAGCGCTTTAAATAGAGGATATGAACCAGATGATTGTTCAAATTGGC 14322

Qy 196 TGCATTGGCTTGCTCAAGGCCGTTGACAAGTTGATCTTCTGACGATGTGAGATTTCG 255  
 Db 14321 TGTATTGGATTGCTAAAATCCGTTGATAAATTGATTGATATGATGTGAAGTTTCG 14262

Qy 256 ACCTATGGGTGCCAATGATCATCGGAGAAATTCAACGCTTTGCGGATGACGGTACG 315  
 Db 14261 ACCTATGCTGTTCAATGATTATTGGTAAATCCAGGATTCACTGGGATGATGGAACA 14202

Qy 316 GTTAAGGTCAGTCGATCGTTAAAGAAACAGCGAATAAGGTGGCGGATCAAAGGATGAA 375  
 Db 14201 GTCAAAGTCAGCGGTCACTAAAGGAACCTGGAAATAAAATCCGGGGCAAGAGATGAA 14142

Qy 376 TTGTACAAGCAATTGGCGGTGCCCCACGATGGCAGAAGTGGCAGAGCAGTGGATATC 435  
 Db 14141 TTGTCTAAGTCTCATGGAAGAATGCCGACCGTCCAGGAAATAGCCGAATATCTAGATATC 14082

Qy 436 ACGCCGGAGGAAGTAGTCCTTGGCAAGAGGCAAGCAGAGGCGCTTCCATCCATGAG 495  
 Db 14081 ACACCAAGGGATGTCGTCCTCGCCCAAGAAGCGGTCCGAACGGCGTGGATTGAA 14022

Qy 496 ACCGTTTGAAATGACGGCGATCCCACACTGATGATCAGATAGCGGATGAAGGT 555  
 Db 14021 ACGGTTATGAAATGATGGGGACCCATTACATTACTCGATCAAATTGCTGATCATTCA 13962

Qy 556 GTGAACAAGTGGTTGAGAAAATTGCTTGAAGGACGCCATCAGCAGGCTGACCGAGCGT 615  
 Db 13961 GAAGAAAGATGGTTGATAAAATCGCGCTGAAAGAAGCGATCAAGGAATTAGAACGA 13902

Qy 616 GAGCAGCTCATCGTCTACCTGCGCTATTACAAGGATCAGACACAGTCTGAGGTAGCAGAG 675  
 Db 13901 GAAAAACTCATTGTGACTTAAGGTATTATAAGACCAACACAATCTGAGGTAGCAGAG 13842

Qy 676 CGTCTAGGGATTGCGAGGTCCAGGTCTCGCGTCTGGAAAGCGTATCCTGCTAACGATC 735  
 Db 13841 AGACTCGGCATTCACAAGTACAAGTCTCGCGCTCGAAAGAAAATTAAAACAAATC 13782

Qy 736 AAGGAGCAAATTGAACA 752  
 Db 13781 CAAATGCAAATGGATCA 13765

## RESULT 8

CP000557\_23/c

## WPCOMMENT

Sequence split into 36 fragments LOCUS CP000557 Accession CP000557

Fragment Name	Begin	End
CP000557_00	1	110000
CP000557_01	100001	210000

CP000557_02	200001	310000
CP000557_03	300001	410000
CP000557_04	400001	510000
CP000557_05	500001	610000
CP000557_06	600001	710000
CP000557_07	700001	810000
CP000557_08	800001	910000
CP000557_09	900001	1010000
CP000557_10	1000001	1110000
CP000557_11	1100001	1210000
CP000557_12	1200001	1310000
CP000557_13	1300001	1410000
CP000557_14	1400001	1510000
CP000557_15	1500001	1610000
CP000557_16	1600001	1710000
CP000557_17	1700001	1810000
CP000557_18	1800001	1910000
CP000557_19	1900001	2010000
CP000557_20	2000001	2110000
CP000557_21	2100001	2210000
CP000557_22	2200001	2310000
CP000557_23	2300001	2410000
CP000557_24	2400001	2510000
CP000557_25	2500001	2610000
CP000557_26	2600001	2710000
CP000557_27	2700001	2810000
CP000557_28	2800001	2910000
CP000557_29	2900001	3010000
CP000557_30	3000001	3110000
CP000557_31	3100001	3210000
CP000557_32	3200001	3310000
CP000557_33	3300001	3410000
CP000557_34	3400001	3510000
CP000557_35	3500001	3550319

Continuation (24 of 36) of CP000557 from base 2300001 (CP000557 Geobacillus thermodeni)

Query Match 46.8% Score 353.6; DB 14; Length 110000;  
Best Local Similarity 68.9% Pred. No. 1.2e-101;  
Matches 485; Conservative 0; Missmatches 219; Indels 0; Gaps 0;

Qy	53	TGAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACGGATGCACGTGAGCTTCTCGTGA 112
Db	58581	TGAAGGAGCTGATCGCGCCAGTCAGGAAGGCAGCCAAGAACCGCGCGATGAAATTATTG 58522
Qy	113	ATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGGGCTTATCAACCGGGGTATGAAG 172
Db	58521	AAAAAAATATGCGCTCGTCTGGTCCGTCCAGGGTTCTTAAATCGTGGCTATGAAG 58462
Qy	173	CGGATGATTTGTTTCAAGATGGTTGCATGGCTTGCCTAAGGGCGTTGACAAGTTGATC 232
Db	58461	CGGATGATTTGTTCAAATCGCTGCATGGCTTGTAAAATCAGTCGATAAGTTGACT 58402
Qy	233	TTTCGTAACGATGTGAGATTTTCAACCTATGCGGTGCCAATGATCATGGAGAAATTCAAC 292
Db	58401	TATCGTATGACGTCAAGTTTCAACGTAACGCTGGCGATGATCATGGGGAGATTCAAC 58342
Qy	293	GCTTTTGCGCGATGACGGTAACGGTTAAGGTCACTCGATCGTAAAGAAACAGCGAATA 352
Db	58341	GATTTCTCCCGATGACGGCACGGTCAAAGTCAGCGCTCGCTGAAAGAGATGGCAATA 58282
Qy	353	AGGTGCGGCGATCAAAGGATGAATTGTACAAGCAATTGGCGGTGCCCCACGATCGCAG 412
Db	58281	AAATCCGCAAAGCGAAGGACGAACGTGTCAGACGGCGGGCGGGCGACGATCACCG 58222
Qy	413	AAAGTGGCAGAAGCAGTGGGAATCAACGGGGAGGAAGTAGTCCTTGCGCAAGAGGGCAAGCCA 472
Db	58221	AAATCGCGGACCATTTGGGTGTTGCGCGGAAGAAGTCGTTCTGCGCAAGAGGGGTTG 58162
Qy	473	GAGOGCCTTCTCCATCCATGAGACCGTTTGTAAATGACGGCGATCCCATCACACTGA 532

Db 58161 GCTCGCGGCGT CGATCCATGAAACAGTGTATGAAAACGACGGCGATCCGATCAAGCTGC 58102  
 Qy 533 TCGATCAGATAGCGGATGAAGGTGTGAACAAGTGGTTGAGAAAATTGCCTTGAAGGACG 592  
 Db 58101 TCGATCAAATCGCCGATGCTGACGAGGCGTATGGTTGAGAAAATTGCCTTAAAAAAAG 58042  
 Qy 593 CCATCACCGAGGCTGAGCGAGCGTGCACAGCTCATCGTCTACCTGCGCTATTACAAGGATC 652  
 Db 58041 CGATTGAAGAGCTCGATGAGCGCGAAAGGCTATCGTCTATTGCGTTATTACAAGGAC 57982  
 Qy 653 AGACACAGTCTGAGGTAGCAGAGCGCTAGGGATTGCGAGGTCCAGGTCTCGCGCTGG 712  
 Db 57981 AAACCCAGTCGAAGTGGCAGCAAGACTGGCATCTCTCAAGTCCAAGTATCCGGCTGG 57922  
 Qy 713 AAAAGCGTATCCTGCTAACGATCAAGGAGCAAATTGAACATTAG 756  
 Db 57921 AAAAGAAAATATTGCAGCACATAAGGAAAAATGGATGGTAG 57878

## RESULT 9

AP006627\_19

WPCOMMENT

Sequence split into 43 fragments LOCUS AP006627 Accession AP006627

Fragment Name	Begin	End
AP006627_00	1	110000
AP006627_01	100001	210000
AP006627_02	200001	310000
AP006627_03	300001	410000
AP006627_04	400001	510000
AP006627_05	500001	610000
AP006627_06	600001	710000
AP006627_07	700001	810000
AP006627_08	800001	910000
AP006627_09	900001	1010000
AP006627_10	1000001	1110000
AP006627_11	1100001	1210000
AP006627_12	1200001	1310000
AP006627_13	1300001	1410000
AP006627_14	1400001	1510000
AP006627_15	1500001	1610000
AP006627_16	1600001	1710000
AP006627_17	1700001	1810000
AP006627_18	1800001	1910000
AP006627_19	1900001	2010000
AP006627_20	2000001	2110000
AP006627_21	2100001	2210000
AP006627_22	2200001	2310000
AP006627_23	2300001	2410000
AP006627_24	2400001	2510000
AP006627_25	2500001	2610000
AP006627_26	2600001	2710000
AP006627_27	2700001	2810000
AP006627_28	2800001	2910000
AP006627_29	2900001	3010000
AP006627_30	3000001	3110000
AP006627_31	3100001	3210000
AP006627_32	3200001	3310000
AP006627_33	3300001	3410000
AP006627_34	3400001	3510000
AP006627_35	3500001	3610000
AP006627_36	3600001	3710000
AP006627_37	3700001	3810000
AP006627_38	3800001	3910000
AP006627_39	3900001	4010000
AP006627_40	4000001	4110000
AP006627_41	4100001	4210000
AP006627_42	4200001	4303871

Continuation (20 of 43) of AP006627 from base 1900001 (AP006627 *Bacillus clausii* KSM-1

Query Match 46.1% Score 348.4; DB 14; Length 110000;  
 Best Local Similarity 68.2% Pred. No. 5.6e-100;  
 Matches 484; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

Qy 37 CTGACCAATGACCAAGTGAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACGGATGCA 96  
 Db 12188 CTATCGATAAACAAAGTGAAGAGCTGATTGCAAAAAGCCAGGAAGGCACACAGAAGCA 12247

Qy 97 CGTGAGCTCTCGTGAATAGCAATATCAGACTGGCTGGTCGTCAGCGCTTATC 156  
 Db 12248 CGGGATTGATCGTCAACCATAACACACAGTCTCGTCTGGTCAGTCGTCAACAGTTTTG 12307

Qy 157 AACCGCCGGTATGAAGCGGATGATTGTTTCAAGATCGGTTGCATTGGCTGCTCAAGGCC 216  
 Db 12308 AATCGCCGTTATGAGGCAGACGACCTTTCAAATTGGCTGCATTGGTTGATTAAAGTCT 12367

Qy 217 GTTGACAAGTTCGATCTTACGATGTGAGATTTGACCTATGCGGTGCGAT 276  
 Db 12368 GTCGACAAATTGACCTTCTATGATGTGAAATTCTCCACGTATGCTGTGCGAT 12427

Qy 277 ATCGGAGAAATTCAACGCTTTTGCACGATGACGGTACGGTTAAGGTCACTCGATCGTTA 336  
 Db 12428 ATTGGGGAAATCCAACGTTTTTGCACGATGATGGTACGGTAAAGTAAGCCGGTCCATT 12487

Qy 337 AAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAATTGTAACAGCAATTGGCGGT 396  
 Db 12488 AAAGAATTAAAGCAATAAAATTGCAAAAGCAAAAGACGAGCTGACGAAACGCTGCGCGG 12547

Qy 397 GCGCCACGATCGCAGAAGTGGCAGAACGAGCTGGAAATCACGCCGGAGGAAGTAGTCTT 456  
 Db 12548 GCGCGGACCATTAATGAGATTGCGAACATTAGCGTGAACGCTGAGGAAATTGTATT 12607

Qy 457 GCGCAAGAGGCAAGCAGAGGCGCTTCCCTCATCCATGAGACCGTTTGGAAATGACGGC 516  
 Db 12608 GCTGGAGATGCCAACCGGAGCTTGTCTCAATCCATGAAACGGTCTATGAAATGATGGC 12667

Qy 517 GATCCCCATCACACTGATCGATCAGATAGGGATGAAGGTGTGAAACAAGTGGTTGAGAAA 576  
 Db 12668 GACCCGATTACACTGCTAGATCAAATTGCGACCACTCACAAAGTCAGTGGTTGACAAG 12727

Qy 577 ATTGCCTTGAAGGACGCCATCAGCAGGCTGAGCGAGCGTGGACAGCTCATCGTCTACCTG 636  
 Db 12728 ATTGCTTTAAAAGAAGCCATTGCGACCTTGGCGAAAGGGAGCGGCTGATAGTATACTTG 12787

Qy 637 CGCTATTACAAGGATCAGACACAGTCTGAGGTAGCAGAGCGCTAGGGATTGCGAGGT 696  
 Db 12788 CGTTATTATAAGGATCAAACACAATCAGAAGTGGCTGAACGCCCTAGGCATTGCAAGTG 12847

Qy 697 CAGGTCTCGCGTCTGGAAAAGCGTACCTGCTAACGATCAAGGAGCAAAT 746  
 Db 12848 CAAGTTCTCGGCTAGAGAAAAAAATTGGAACAAATGAAAGAAACAAT 12897

## RESULT 10

BACSP02A

LOCUS BACSP02A 864 bp DNA linear BCT 26-APR-1993

DEFINITION B. subtilis spolI A gene encoding a sporulation-specific sigma factor, complete cds.

ACCESSION M15744

VERSION M15744.1 GI:143610

KEYWORDS sporulation-specific sigma factor.

SOURCE Bacillus subtilis

ORGANISM Bacillus subtilis

Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

REFERENCE 1 (bases 1 to 864)

AUTHORS Yudkin, M D.

TITLE Structure and function in a Bacillus subtilis sporulation-specific sigma factor: molecular nature of mutations in spolIAC

**JOURNAL** J. Gen. Microbiol. 133 (Pt 3), 475-481 (1987)  
**PUBMED** 3116160  
**COMMENT** Original source text: *B. subtilis* (strain 168) DNA, clone BgIII/PstI.  
 Draft entry and clean copy sequence for [1] kindly provided by M D. Yukin, 26-MAY-1987.  
**FEATURES** Location/Qualifiers  
**source**  
 1..864  
 /organism="Bacillus subtilis"  
 /mol\_type="genomic DNA"  
 /strain="W168"  
 /sub\_strain="PY79"  
 /db\_xref="taxon:1423"  
**CDS**  
 90..857  
 /note="sporulation-specific sigma factor"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="AAA22788.1"  
 /db\_xref="GI:143611"  
 /translation="MDVEVKKNGKNAQLKDHEVKELIKOSQNGDQOARDLLI EKNMRL  
 VWSVVQRFLNRGYEPDDL FQI GCI GLLKSVDKFDLTYDVRFSTYAVPM I GEI QRFI R  
 DDGTVKVSRSLKELGNKI RRAKDEL SKT LGRVPTVQEI ADHLEI EAEDVWLAQEAVRA  
 PSSI HETVYENDGDPITLLDQI ADNSEEKWFDFKI ALKEAI SDLEEREKLI VYLRYYYKD  
 QTQSEVAERLGISQVQVSRLKKILKQIKVQMDHTDG"  
**ORIGIN** 2 bp upstream of *Cla*I site.  
 Query Match 45.7% Score 345.6; DB 14; Length 864;  
 Best Local Similarity 66.5% Pred. No. 2. 1e-99;  
 Matches 495; Conservative 0; Mismatches 249; Indels 0; Gaps 0;  

Qy	10	GATATCAAAAATGCGAGTCACCAATTCTGACCAATGACCAAGTGAAAGATTTGATAGCC	69
Db	102	GTTAAGAAAAACGGCAAAAACGCTCAGCTGAAGGATCATGAAGTAAAGAATTAAATCAAA	161
Qy	70	AAGAGCCAAGCTGGCGATAOGGATCCACGTGAGCTTCTCGTGAATAGCAATATCAGACTG	129
Db	162	CAAAGCCAATGGCGACCAGCAGCCAAAGAGACCTCCTCATAGAAAAAAACATGGTCTT	221
Qy	130	GTCTGGTCGTCGTOCAGOGCTTATCAACCGGGGTATGAAGGGATGATTGTTTCAG	189
Db	222	GTTGGTCTGTCGTACAGOGGTTAAACAGAGGATATGACCTGACGATCTCTCCAG	281
Qy	190	ATCGGTTGCATTGGCTTGCTCAAGGCCGTTGACAAGTTGATCTTGTACGATGTGAGA	249
Db	282	ATCGGCTGCATCGGGCTGTTAAACAGAGGATATGACCTGACGATCTCTCCAG	341
Qy	250	TTTTCGACCTATGGGTGCGAATGATCATCGGAGAAATTCAACGTTTGGCGATGAC	309
Db	342	TTTTCAACGTATGCAGTGCAGATGATTATCGGAGAAATCCAACGTTATCGTGTGAC	401
Qy	310	GGTACCGTTAAGGTCACTGGTCAAGTGGTAAAGGAAACAGCGAATAAGGTGGCGATCAAAG	369
Db	402	GGAACCGTAAAGGTATCACGGTCATTAAAAGAGCTTGGAAACAAAATCCGGCGCGAAG	461
Qy	370	GATGAATTGTACAAGCAATTGGCCGTGCCCCACGATGCGAGAAGTGGCAGAAGCAGTG	429
Db	462	GATGAGCTTCGAAAACACTGGGAGAGTGGCGACGGTGCAGGAGATCGCTGACCATTTG	521
Qy	430	GGAATCACGCCGGAGGAAGTAGTCCTTGCGCAAGAGGCAAGCAGAGGGCTTCCATC	489
Db	522	GAGATTGAAGCTGAGGATGTTGACTGGCCCAAGAGGGCTTAAAGGCTCCATCTCGATT	581
Qy	490	CATGAGACCGTTTGAAATGACGGCGATCCATCACACTGATCGATCAGATAGGGAT	549
Db	582	CACGAAACCGTTATGAAATGACGGAGATCCGATTACCGTGCCTGATCAAATCGCTGAC	641
Qy	550	GAAGGTGTGAACAAGTGGTTGAGAAAATTGCCTTGAAGGACGCCATCAGCAGGCTGAGC	609
Db	642	AACTCAGAAGAAAATGGTTGACAAAATTGCGCTGAAAGAAGGATCAGCGATTGGAG	701

Qy 610 GAGCGTGAGCAGCTCATCGTCTACCTGCGTATTACAAGGATCAGACACAGCTGAGGT A 669  
 Db 702 GAAAGGGAAAAACTAATCGTCTATCTCAGATATTATAAGACCAACAGTCGGAGGTG 761  
 Qy 670 GCAGAGCGTCTAGGGATTGCGAGGTCAGGCTCGCGTCTGGAAAAGCGTATCCTGCTA 729  
 Db 762 GCTGAGCGGCTCGGGATCTCAGGTCAGGTTCCAGGTTGAAAAGAAAATTAAAAA 821  
 Qy 730 ACGATCAAGGAGCAAATTGAACAT 753  
 Db 822 CAGATCAAGGTTCAAATGGATCAT 845

## RESULT 11

Z99116/ c

**LOCUS** Z99116 204537 bp DNA linear BCT 18- APR- 2005  
**DEFINITION** *Bacillus subtilis* complete genome (section 13 of 21): from 2409151 to 2613687.  
**ACCESSION** Z99116 AL009126  
**VERSION** Z99116.2 G : 32468787  
**KEYWORDS**  
**SOURCE** *Bacillus subtilis* subsp. *subtilis* str. 168  
**ORGANISM** *Bacillus subtilis* subsp. *subtilis* str. 168  
**BACTERIA; Firmicutes; Bacillales; Bacillaceae; Bacillus.**  
**REFERENCE** 1 (bases 1 to 204537)  
**AUTHORS** Kunst, F., Ogasawara, N., Mbszer, I., Albertini, A. M., Altoni, G., Azevedo, V., Berto, M. G., Bessieres, P., Bolotin, A., Borchart, S., Borri, R., Bourssi, L., Brans, A., Braun, M., Brignell, S. C., Bron, S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., Carter, N. M., Choi, S. K., Codani, J. J., Connerton, I. F., Cummings, N. J., Daniel, R. A., Denizot, F., Devine, K. M., Dusterhoff, A., Ehrlisch, S. D., Emmerson, P. T., Entian, K. D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S. Y., Glaser, P., Goffeau, A., Golightly, E. J., Grandi, G., Giuseppe, G., Guy, B. J., Haga, K., Haech, J., Harwood, C. R., Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M. F., Itaya, M., Jones, L., Joris, B., Karamat, A., Kasahara, Y., Klaerr-Blanchard, M., Klein, C., Kobayashi, Y., Koettner, P., Konigstein, G., Krugh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarovic, V., Lee, S. M., Levine, A., Liu, H., Masuda, S., Mael, C., Medigue, C., Medina, N., Meliado, R. P., Mizuno, M., Mestel, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S. H., Parro, V., Pohl, T. M., Portetelle, D., Porwollik, S., Prescott, A. M., Presecan, E., Pujic, P., Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadai, Y., Satoh, T., Scanlan, E., Schleicher, S., Schreuter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S. J., Serror, P., Shin, B. S., Soldo, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandebol, M., Vanner, F., Vassarotti, A., Viari, A., Wambutt, R., Wedler, E., Wedler, H., Wittenberger, T., Winters, P., Wpat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshiida, K., Yoshioka, H. F., Zumstein, E., Yoshioka, H. and Danchin, A.  
**TITLE** The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*  
**JOURNAL** Nature 390 (6657), 249-256 (1997)  
**PUBMED** 9384377  
**REFERENCE** 2 (bases 1 to 204537)  
**AUTHORS** Kunst, F., Ogasawara, N., Yoshioka, H. and Danchin, A.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (27-JUN-2003) I. Mbszer, A. Danchin, Institut Pasteur, Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail: mbszer@pasteur.fr, adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45 68 89 48  
**COMMENT** On Jul 7, 2003 this sequence version replaced gi:2634723.

This entry contains data from release R16.1 of the SubtiList database. Further data on gene annotation and detailed information about changes from previous releases can be found at <http://genolist.pasteur.fr/SubtiList/>.

FEATURES	source	Locati on/Qualifiers
		1. . 204537
		/organism="Bacillus_subtilis subsp. subtilis str. 168"
		/mol_type="genomic DNA"
		/strain="168"
		/db_xref="taxon:224308"
gene		complement(61. . 633)
		/gene="ypaA"
		/locus_tag="BSU23050"
CDS		complement(61. . 633)
		/gene="ypaA"
		/locus_tag="BSU23050"
		/function="unknown"
		/codon_start=1
		/transl_table=11
		/protein_id="CAB14237.1"
		/db_xref="GI:2634740"
		/db_xref="GOA: P50726"
		/db_xref="SubtiList: BG11428"
		/db_xref="UniProtKB/Swiss-Prot: P50726"
		/translation="MKVKKLVVVSMSSIAFVLMLLNFPFPGLPDYLKI DFSDVPAI I ALI YGPLAGI AVEAI KNVLQYI I QGSMAGVPVGQVANFI AGTLFI LPTAFLFKKLNS AKGLAVSLLLGTAAMTI LM61 LNYVLI LPAYTWFLHSPALSALSKTAWAGI LPFNM I KGIVI TVVFSLI FI KLPKW EQQRSAHI H"
gene		complement(739. . 903)
		/gene="ypzE"
		/locus_tag="BSU23060"
CDS		complement(739. . 903)
		/gene="ypzE"
		/locus_tag="BSU23060"
		/function="unknown"
		/codon_start=1
		/transl_table=11
		/protein_id="CAB14238.1"
		/db_xref="GI:2634741"
		/db_xref="SubtiList: BG13767"
		/db_xref="UniProtKB/Swiss-Prot: Q82010"
		/translation="MKFPTGGDEPMALSPRAVFTAGFGEI PEPTVOSGWEKMEVHKRF EI EFKRFFA"
gene		1130. . 2707
		/gene="ser A"
		/locus_tag="BSU23070"
CDS		1130. . 2707
		/gene="ser A"
		/locus_tag="BSU23070"
		/EC_number="1.1.1.95"
		/function="serine biosynthesis"
		/codon_start=1
		/transl_table=11
		/product="phosphoglycerate dehydrogenase"
		/protein_id="CAB14239.1"
		/db_xref="GI:2634742"
		/db_xref="GOA: P35136"
		/db_xref="Inserter Protein: PR002912"
		/db_xref="Inserter Protein: PR006139"
		/db_xref="Inserter Protein: PR006140"
		/db_xref="Inserter Protein: PR006236"
		/db_xref="SubtiList: BG10509"
		/db_xref="UniProtKB/Swiss-Prot: P35136"
		/translation="MFRVLVSDKMSNDGLQPLI ESDFI EI VQKNVADAEDELHTFDAL LVRSATKVTEDLFNKMTSLKI VGRAGVGVDNI DI DEATHKGVI VI NAPNGNTI STAEH TFAM SSLMRHI PQANI SVKSREWRAYVGSELYGKTLGI VGLGRI GSEI AQRRGAF GMTVHVFDPLTEERAKKI GVNSRTFEEVLESADI I TVHTPLTKETKGLLNKTETI AKT KKGVRLI NCARGGI DEAALLFALENGHVAGAALDVFEVEPPVDNKLVDHPLVI ATPH LGASTKEAQLNVAACQSEEVLQFAKGLPVMSAI NLPAMTKDEFAKI KPYHQI AGKI GS

LVSQCMKEPVQDVAI QYEGTI AKLETSFI TKALLSGFLKPRVDSTVNEVNAGGVAKER  
 GI SFSEKI SSSESGYDNCI SVKVTGDRSTFTVTATYI PHFGERI VEI NGFNI DFYPTG  
 HLVYI QHQDTTGVII GRVGRI LGDNDI NI ATMQVGRKEKGGEAI MMLSFDRLLEDKI VK  
 ELTNVPDI VSVKLI DLP"

gene compl element (2750. . 3517)  
 / gene="ar oC"  
 / locus\_tag="BSU23080"

CDS compl element (2750. . 3517)  
 / gene="ar oC"  
 / locus\_tag="BSU23080"  
 / EC\_number="4.2.1.10"  
 / function="shikimate pathway"  
 / codon\_start=1  
 / transl\_table=11  
 / product="3-dehydrorquinic acid dehydratase"  
 / protein\_id="CAB14240.1"  
 / db\_xref="GI:2634743"  
 / db\_xref="GOA: P35146"  
 / db\_xref="InterPro: IPR001381"  
 / db\_xref="Subtilisins: BG10538"  
 / db\_xref="UniprotKB/Swiss-Prot: P35146"  
 / translation="MNVLTIGVSI GEGMPKIIPLMGKTEKQLNEAEAVKLLNPDI  
 VEWRVDVFEKANDREAVTKLISKLRKSLEDKLFIFTFRTHKEGGSMEMDESSYLA  
 SAI QTKDI DLI DI ELFSGDANVKALVSLAEENNYYVMSNHDFEKTPVKDEI  
 ISRLRK MQDLGAHI PKMAVMPNDTGDLTLLDATYTMKTYADRPIITMSMAATGLISRLSGEV  
 FGSACTFGAGEEASAPGQIPVSELRSVLDILHKNTRG"

gene compl element (3629. . 4735)  
 / gene="rsi X"  
 / locus\_tag="BSU23090"

CDS compl element (3629. . 4735)  
 / gene="rsi X"  
 / locus\_tag="BSU23090"  
 / function="negative regulation of sigma-X activity"  
 / note="alternative gene name: ypuN"  
 / codon\_start=1  
 / transl\_table=11  
 / protein\_id="CAB14241.1"  
 / db\_xref="GI:2634744"  
 / db\_xref="GOA: P35166"  
 / db\_xref="Subtilisins: BG10537"  
 / db\_xref="UniprotKB/Swiss-Prot: P35166"  
 / translation="MMKSEWNEEQIKELLSQLPAVKDHRSPOQIYKRLTMAKRKNKPA  
 VRWGPACAAAI AVYIAFIISPHFFDQACPQQKEASQENAVTKTETEDSPKAASSLDQ  
 TSFVVEKEQDNYITVAVADADTSAIIPVSIQKTNADQTIQDMLFESSELGLDHAIT  
 IPTFI DEVEI KEKPKQKELSI RVHQPATAFSI KDDTLLKKL LKESLKWSPYEKVKFLS  
 DQNETGVRI GSYGTFTEISIPKQSKRSYYLYQNKGQDFL VPSNHSFDTVKEAI KEME  
 SSSQEDTTPLI QAGAVQSVTKKQKHL YMRFSKESEVDDSIAGILM EGLLTAKEFGF  
 TEVTFTETRTKKI GKYDI SDAIPVPAAPNPI SLN"

gene compl element (4671. . 5255)  
 / gene="sigX"  
 / locus\_tag="BSU23100"

CDS compl element (4671. . 5255)  
 / gene="sigX"  
 / locus\_tag="BSU23100"  
 / function="may participate in the regulation of  
 peptidoglycan synthesis and turnover"  
 / note="alternative gene name: ypuM"  
 / codon\_start=1  
 / transl\_table=11  
 / product="RNA polymerase EC( extracytoplasmic  
 function)-type sigma factor (sigma-X)"  
 / protein\_id="CAB14242.1"  
 / db\_xref="GI:2634745"  
 / db\_xref="GOA: P35165"  
 / db\_xref="InterPro: IPR000838"  
 / db\_xref="InterPro: IPR007627"  
 / db\_xref="InterPro: IPR007630"  
 / db\_xref="InterPro: IPR013324"  
 / db\_xref="InterPro: IPR013325"

/ db\_xr ef ="Subt i Li st : BG10536"  
 / db\_xr ef ="Uni Pr ot KB/ Swi ss- Pr ot : P35165"  
 / t r ans l at i on="MEETFQLLYDTYHQDL YQFL FYM/KDKNQTEDLL QEVYI RVLNS  
 YHTFEGRSSEKTVLLSI ARHVAI DWFRKQQTI RQRI LGTFDWDT QDVRDQQL PDELA  
 VQHENVREI SCALDQCTI DQRAVI I LRFI QGYSI QETAKALRFSESKVKTQHRLKV  
 LRKHMLLREELMDDEVRMERRTDKGWVKSTSGS"  
 gene compl ement ( 5459. . 7228)  
 / gene="resE"  
 / l ocus\_t ag="BSU23110"

Query Match 45.7% Score 345.6; DB 14; Length 204537;  
 Best Local Similarity 66.5% Pred. No. 4.9e-99;  
 Matches 495; Conservative 0; M smatches 249; Indels 0; Gaps 0;

Qy	10	GATATCAAAAATGCGAGTCACCATTCTGACCAATGACCAAGTGAAAGATTGATGCC 69
Db	34223	GTAAAGAAAAACGGCAAAACGCTCAGCTGAAGGATCATGAAGTAAAGGAATTAACTAAA 34164
Qy	70	AAGAGCCAAGCTGGCGATAACGGATGCCAGTGGCTTCCTCGTAATAGCAATATCAGACTG 129
Db	34163	CAAAGCCAAAATGGCGACCAGCAGCCAAGAGACCTCCTCATAGAAAAAAACATGCGCTT 34104
Qy	130	GTCTGGTCGGTCCAGOGCTTTATCAACCGGGGTATGAAGCGGATGATTGTTTAG 189
Db	34103	GTGGTCTGTCGTACAGGGTTAAACAGAGGATATGAGCCTGACGATCTCTCCAG 34044
Qy	190	ATCGGTTGCATTGGCTTGCTCAAGGCGTTGACAAGTTGATCTTGTACGATGTGAGA 249
Db	34043	ATCGGCTGCATCGGCTGTTAAATCTGTTGACAAATTGATTTAACCTATGATGTGCGT 33984
Qy	250	TTTCGACCTATGGGTGCGCAATGATCATCGGAGAAATTCAACGGTTTGGCGATGAC 309
Db	33983	TTTCAACGTATGCAGTGGATGATTATCGGAGAAATCCAACGATTATCGTGTGAC 33924
Qy	310	GGTACGGTTAAGGTCACTGGTAAAGGAAATAGGTGGGGGATCAAAG 369
Db	33923	GGAACCGTAAAGGTATCAAGGTCAATTAAAGAGCTGGAAACAAAATCCGGCCCCGCGAAG 33864
Qy	370	GATGAATTGTACAAGCAATTGGCGGTGCGCCACGATCGAGAAGTGGCAGAACAGTG 429
Db	33863	GATGAGCTTCGAAAACACTGGCAGAGTGGCGACGGTGCAGGAGATCGCTGACCAATTG 33804
Qy	430	GGAATCACGCGGGAGGAAGTAGTCCTTGCGCAAGAGGCAAGCAGAGGGCTTCCATC 489
Db	33803	GAGATTGAAGCTGAGGATGTTGACTGGCCAAGAGGGGTAAGGGCTCCATCTTGATT 33744
Qy	490	CATGAGACCGTTTGAAATGACGGCGATCCCATCACACTGATCGATCAGATAGGGAT 549
Db	33743	CACGAAACCGTTATGAAAATGACGGAGATCGATTACCGTGTGATCAAATCGCTGAC 33684
Qy	550	GAAGGTGTGAAACAAGTGGTTGAGAAAATTGGCTTGAGGGACGCGATCAGCAGGCTGAGC 609
Db	33683	AACTCAGAAGAAAAATGGTTGACAAAATTGGCTGAAAGAAGCGATCAGCGATTGGAG 33624
Qy	610	GAGCGTGAGCAGCTCATCGTCTACCTGCGCTATTACAAGGATCAGACACAGTCTGAGGTA 669
Db	33623	GAAAGGGAAAAACTAATCGTCTATCTCAGATATTATAAAGACCGACACAGTGGAGGTG 33564
Qy	670	GCAGAGCGTCTAGGGATTGCGAGGTCCAGGTCTCGCGTCTGGAAAAGCGTATCCTGCTA 729
Db	33563	GCTGACCGGCTCGGATCTCTCAGGTGCAGGTTCCAGGCTTGAAAAGAAAATTAAAAA 33504
Qy	730	ACGATCAAGGAGCAAATTGAACAT 753
Db	33503	CAGATCAAGGTTCAAATGGATCAT 33480

RESULT 12  
BACJH642



CDS	LGTTLASDEVDTI GGFMTQQI DAAVGSI EADGYI FKVHETVGRHI NYLEI VRKKE" compl ement ( 811. . 1269) / codon_st art =1 / transl_t abl e=11 / product =" Bl t D" / prot ei n_i d=" BAA12354. 1" / db_xref =" GI : 1303698" / transl at i on=" MSI NI KAVTDDNRAAI LDLHVSQNQLSYI ESTKVCLEDAKECHY YKPVGL YYEGDL VGFAMYGL FPEYDEDNKNGRVWL DRFFI DERYQGKGL GKKML KALI QHLAEL YKCKRI YLSI FENNI HAI RLYQRFGFQFNGEL DFNGEKVM/KEL" compl ement ( 1278. . 1284) compl ement ( 1438. . 2640) / codon_st art =1 / transl_t abl e=11 / product =" Bl t " / prot ei n_i d=" BAA12355. 1" / db_xref =" GI : 1303699" / transl at i on=" MKKSI NEQKTI FI I LLSNI FVAFLGI GLI I PVMPSFMKI MHL SG STMGYLVAAFAI SQLI TSPFAGRWDRFGRKMM I LGGLI FSLSELI FGLGTHVSI FY FSRI LGGVSAAFI MPAVTAYVADI TTLKERSKAMGYVSAAI STGFI I GPGAGGF1 AGF GI RMPFFFASAI ALI AAVTSVFI LKESLSI EERHQLSSHTKESNFI KDLKRSI HPVYF I AFI I VFVMAFGL SAYETVFSLFSDHKFGFTP KDI AAI I TI SSI VAVVI QVLLFGKL V NKLGEKRM QLCI TGAI LAFVSTVMSGFLTVL VTCFI FLAFDLLRPALTAHLSNMA GNQQGFVAGVNSTYSLGNI FGPA LGGI LF DNL NI HYPFLFAGFVM VGLGLTMWKEK KNDAALAN"
RBS	compl ement ( 2648. . 2654)
RBS	2745. . 2752
CDS	2757. . 3578 / codon_st art =1 / transl_t abl e=11 / product =" Bl t R" / prot ei n_i d=" BAA12356. 1" / db_xref =" GI : 1303700" / transl at i on=" MSEDVKKYFTTGEFSKL CRVKKQTL FHYDEI GLFSPEI KKENGY RYYSYHQFETFQVI SLFKELGVPLKEI KCLI KGKTPDKI LHLVKEKSI EI DKKI NELK QLQTI LQTKVTLTEQAL ETD FSSI SFEYNEETFMLS RSKTL NLPERKYVAAI SELI HE VQQYEL DEGYPI GGI FAREQI LEKDFYNYSYFYI KVKGDAENI NYHVRPKGL YAVGYE I GGNTEEAYRRI I EFi ERNGMqi GENAYEEYML DEMVVDGYENTYAKI LLQVKEV"
RBS	3731. . 3736
CDS	3748. . 3894 / codon_st art =1 / transl_t abl e=11 / product =" YrkB" / prot ei n_i d=" BAA12357. 1" / db_xref =" GI : 1303701" / transl at i on=" MLLKNWPSRRI QRDKSKRAGI GGTNNRI PYTLLL CYVN VQKPR I VDL"
RBS	4108. . 4111
CDS	4120. . 4680 / codon_st art =1 / transl_t abl e=11 / product =" YrkC" / prot ei n_i d=" BAA12358. 1" / db_xref =" GI : 1303702" / transl at i on=" MYYNPFSPOCYYYVTVPMYNDGRSVYWTI PNEMEKVHRGADLRS SYEDRNILLKDYGPKPFWNI NRATKQNTFRTA WTGHFQVTLMSLG GEDI GLEI HPNVDQFLRI EQGRGI VKMGKSKDHLN FQRNVYDDSAI WPAGTWHNVI NTGNTPLKL YSI YAPPNHPFGTVHETKADAVAAED"
RBS	5458. . 5464
CDS	5473. . 5664 / codon_st art =1 / transl_t abl e=11 / product =" YrkD" / prot ei n_i d=" BAA12359. 1" / db_xref =" GI : 1303703" / transl at i on=" MVEQQKDCREVVTQLAASRNAI DRAMGLI VSTNL EHCVRESLEK GEDTQNLVKEAVDVLVKS"
RBS	5797. . 5801
CDS	5809. . 6291

RBS	CDS	/codon_start=1 /transl_table=11 /product="YrkE" /protein_id="BAA12360. 1" /db_xref="GI : 1303704" /transl_aton="MTEQTAKRTTI VLFSGDYDKAMAAYI I ANGAAAYDHEVTI FHTFW GFNALRKEELI PVKKGFLAKMFGKMPRGADKMLSKMNFAGVGPKM KNVMKKHNVL TLPQLI EMAQEQQV/KL VACTMTMDLLGLQEKELLDDI DYAGVAAYLADAEGSVNLF 6464. . 6471 6477. . 7034
RBS	CDS	/codon_start=1 /transl_table=11 /product="YrkF" /protein_id="BAA12361. 1" /db_xref="GI : 1303705" /transl_aton="MMKATI VLDAGKLAGCPMPI VKTKKRMKDLKAGEVLEI HATDKGS TADLEAWAKSTGHEYLGTEAEGEI L RHFLRKGGEHSSENASSI PEI SLEAFKQKVDSL ESLNI LDVREI EYEYKAHI PGVWHI PLGEVEKRNANELNENDEI YI I CHSGRRSEMAAR TMKKQGFKKVI NVVPGMRDWTGKTE" 7361. . 7367 7376. . 7486
RBS	CDS	/codon_start=1 /transl_table=11 /product="YrkG" /protein_id="BAA12362. 1" /db_xref="GI : 1303706" /transl_aton="MTVKAMTPKEVTKSLTKNLYLFWWCVMKM LVI GR" 7597. . 7605 7615. . 8502
RBS	CDS	/codon_start=1 /transl_table=11 /product="YrkH" /protein_id="BAA12363. 1" /db_xref="GI : 1303707" /transl_aton="MVSEMLSEAGHSVSYLGKGMKAWSEHLEPVKVGDLDKGGEEMYQF VRI GKGCLSYMMI SNGEAAI DATRMNFFIDFAKGKGATI THVFDTHLHADHISGGR VI AEKTAKTYLPPKDAEEVEFDYQPLEEGNDI I I GNNTI KI QPI YSPGHTI GSTSF VDDQYLLSGDI LFI DSI GRPDLAGMAEDWGDLRKTLYSRYKALSKDLI VLPAHFM I DELNENGSGVKQLSTLFSENHGLNI ADETEFRHLVTDHLPPQPNAYQEI RETNMGKI S PDEEKQREMEI GPNRCAI R" 8523. . 8527 8536. . 8763
RBS	CDS	/codon_start=1 /transl_table=11 /product="YrkI" /protein_id="BAA12364. 1" /db_xref="GI : 1303708" /transl_aton="MKSDFVLDAGKLAGCPMPI VRTKKAMNELESGQI LEVHATDKGAK NDLTAWSKSGGHDLLEQTDEGDI LKFW QKG" 8807. . 8815 8824. . 9609
RBS	CDS	/codon_start=1 /transl_table=11 /product="YrkJ" /protein_id="BAA12365. 1" /db_xref="GI : 1303709" /transl_aton="MDIAFI TTLFI I GFI GAFI SGMI GIGGAVI NYPMLLYI PSLVGV MSLTAHEVSGI GAI QVFFATLGGVWYRKSGLI NKTLLI YMGSII LLGSVLSGSYFSHY I SEKGII NFI YGI LAI I AVI LMFI PKKGQKGQEHSEDKEVI FNKWLASSLAFI I GGVSG I LGAGGAFI LVPI MLSI LNI PVRVTVASSLAI TFLSSI GATVGKVI TGQVLFVPALVL M TSLLI ASPI GASVGQKVNTKFLQWLAQLI GATAIKI WLELL" 10177. . 10182

Query Match 45.7% Score 345.6; DB 14; Length 282700;  
Best Local Similarity 66.5% Pred. No. 5.2e-99;  
Matches 495; Conservative 0; Mismatches 249; Indels 0; Caps 0;

Qy 10 GATATCAAAATGCGAGTCACCCATTCTGACCAATGACCAAGTGAAAGATTGATAGCC 69

Db 275466 GTTAAGAAAAACGGCAAAACGCTCAGCTGAAGGATCATGAAGTAAAGGAATTAATCAA 275525  
 Qy 70 AAGAGCCAAAGCTGGCGATAACGGATCCACGTGAGCTTCTCGTGAATAGCAATATCAGACTG 129  
 Db 275526 CAAAGCCAAAATGGCGACCAGCAGCAAGAGACCTCCTCATAGAAAAAAACATGCGTCTT 275585  
 Qy 130 GTCCTGGTCTCGTCCAGCGCTTTATCAACCGCGGGTATGAAGCGGATGATTTGTTTCAG 189  
 Db 275586 GTTTGGTCTGTCGTACAGCGGTTTTAACAGAGGATATGACCTGACGATCTCTCCAG 275645  
 Qy 190 ATCGGTTGCATTGGCTTGCTCAAGGCCGTTGACAAGTTCGATCTTCGATACGATGTGAGA 249  
 Db 275646 ATCGGCTGCATCGGGCTGTTAAAATCTGTTGACAAATTGATTTAACCTATGATGTGCGT 275705  
 Qy 250 TTTTCGACCTATGCGGTGCGAATGATCATCGGAGAAATTCAACGCTTTTGCCTGATGAC 309  
 Db 275706 TTTTCAACGTATGCACTGCGATGATTATCGGAGAAATCCAACGATTTATCCGTGATGAC 275765  
 Qy 310 GGTACCGTTAAGGTCACTCGATCGTTAAAAGAAACAGCGAATAAGGTGCGCGATCAAAG 369  
 Db 275766 GGAACCGTAAAGGTATCAACGCTTAAAGAGCTTGGAAACAAAATCCGGCGCGCGAAG 275825  
 Qy 370 GATGAATTGTACAAGCAATTGGCCGTGCCCCCACGATCGCAGAAGTGGCAGAAGCAGTG 429  
 Db 275826 GATGAGCTTCGAAAACACTGGGAGAGTGGCGACGGTGCAGGAGATCGCTGACCATTTG 275885  
 Qy 430 GGAATCACGCGGGAGGAAGTAGTCTTGGCAAGAGGCAAGCAGAGGCCCTTCCCTCATC 489  
 Db 275886 GAGATTGAAGCTGAGGATGTTGACTGGCCAAGAGGGCGTAAGGGCTCCATCTTGATT 275945  
 Qy 490 CATGAGACCGTTTGAAATGACCGCGATCCCATCACACTGATCGATCAGATAGCGGAT 549  
 Db 275946 CACGAAACCGTTATGAAAATGACGGAGATCGGATTACCGTGCCTGATCAAATCGCTGAC 276005  
 Qy 550 GAAGGTGTGAACAAGTGGTTGAGAAAATTGGCTTGAAGGACGCCATCAGCAGGCTGAGC 609  
 Db 276006 AACTCAGAAGAAAATGGTTGACAAAATTGCGCTGAAAGAAGCGATCAGCGATTGGAG 276065  
 Qy 610 GAGCGTGAGCAGCTCATCGTCTACCTGCGTATTACAAGGATCAGACACAGCTGAGGT 669  
 Db 276066 GAAAGGGAAAACAACTATCGTCTATCTCAGATTATAAGACCAAGACACAGTCCGAGGTG 276125  
 Qy 670 GCAGAGCGTCTAGGGATTTOGCAGGTCCAGGTCTCGCGTCTGGAAAAGCGTATCCTGCTA 729  
 Db 276126 GCTGAGCGGCTCGGGATCTCTCAGGTGCAGGTTCCAGGCTGAAAAGAAAATTAAAAA 276185  
 Qy 730 ACGATCAAGGAGCAAATTGAACAT 753  
 Db 276186 CAGATCAAGGTTCAAATGGATCAT 276209

## RESULT 13

BA000004\_16

WPCOMMENT

Sequence split into 42 fragments LOCUS BA000004 Accession BA000004

Fragment	Name	Begin	End
BA000004_00		1	110000
BA000004_01		100001	210000
BA000004_02		200001	310000
BA000004_03		300001	410000
BA000004_04		400001	510000
BA000004_05		500001	610000
BA000004_06		600001	710000
BA000004_07		700001	810000
BA000004_08		800001	910000
BA000004_09		900001	1010000
BA000004_10		1000001	1110000
BA000004_11		1100001	1210000
BA000004_12		1200001	1310000

BA000004_13	1300001	1410000
BA000004_14	1400001	1510000
BA000004_15	1500001	1610000
BA000004_16	1600001	1710000
BA000004_17	1700001	1810000
BA000004_18	1800001	1910000
BA000004_19	1900001	2010000
BA000004_20	2000001	2110000
BA000004_21	2100001	2210000
BA000004_22	2200001	2310000
BA000004_23	2300001	2410000
BA000004_24	2400001	2510000
BA000004_25	2500001	2610000
BA000004_26	2600001	2710000
BA000004_27	2700001	2810000
BA000004_28	2800001	2910000
BA000004_29	2900001	3010000
BA000004_30	3000001	3110000
BA000004_31	3100001	3210000
BA000004_32	3200001	3310000
BA000004_33	3300001	3410000
BA000004_34	3400001	3510000
BA000004_35	3500001	3610000
BA000004_36	3600001	3710000
BA000004_37	3700001	3810000
BA000004_38	3800001	3910000
BA000004_39	3900001	4010000
BA000004_40	4000001	4110000
BA000004_41	4100001	4202352

Continuation (17 of 42) of BA000004 from base 1600001 (BA000004 Bacillus halodurans C

Query Match 45.4% Score 343; DB 14; Length 110000;  
 Best Local Similarity 67.3% Pred. No. 3. 1e-98;  
 Matches 484; Conservative 0; Missmatches 235; Indels 0; Gaps 0;

Qy	38	TGACCAATGACCAAGTGAAGAGATTGATAGCCAAGAGCCAAGCTGGCGATAACGGATGCAC 97
Db	14367	TTACAGACGGAGAAATTAAAGCAGCTATTGGCAAAGTCAGGACGGAGACCAAGAGGGCA 14426
Qy	98	GTGAGCTTCTCGTGAATAGCAATATCAGACTGGCTGGTCCGTGTCAGCGCTTATCA 157
Db	14427	GAGATCAAATCGTCAGTCGAACACTCGCCTCGTATGGCTGTGCAACGATTCTAA 14486
Qy	158	ACCGCGGGTATGAAGCGGATGATTGTTCAAGATCGGTTGCATTGGCTGCTCAAGGCCG 217
Db	14487	ATCGCGGGTATGAAGCAGACGACCTGTTCAAATTGGTGCATTGGCTGATTAAATCAG 14546
Qy	218	TTGACAAGTCGATCTTCGTACGATGTGAGATTTCGACCTATGCGGTGCCAATGATCA 277
Db	14547	TGGATAAATTGATTGTCGTACGATGTCAAGTTCCACCTATGCGGTGCCAATGATCA 14606
Qy	278	TCGGAGAAATTCAACGCTTTGCGCGATGACGGTACGGTTAAGGTCAAGTCAGTCGATCGTTAA 337
Db	14607	TCGGTGAATCCAACGATTCCCGATGATGGTACGGTAAGGTAAAGTCGCTCCATTAA 14666
Qy	338	AAGAAACAGCGAATAAGGTGCGCGATCAAAGGTGAATTGACAAGCAATTGGCGTG 397
Db	14667	AGGAGCTTGGCAATAAAATTGGAGCTCAAGGTGATGACGAAGACGCTGGTCGGA 14726
Qy	398	CCCCACGATCGCAGAAGTGGCAGAACGAGCTGGGAATCACGCCGGAGGAAGTAGTC 457
Db	14727	CGCCGACGGTGAATGAAATTGGGAGCAGTTAGAAATTACGCCAGAAGAAGTCGTC 14786
Qy	458	CGCAAGAGCCAAGCAGAGCGCCTTCCTCCATCCATGAGACCGTTTGAAATGACGGCG 517
Db	14787	CGGGGGATGCAAGTCGCTCGCTAAGTCCATCCATGAGACTGTCTATGAAATGACGGCG 14846
Qy	518	ATCCCCATCACACTGATCGATCAGATAGCGGATGAAGGTGTGAACAAGTGGTTGAGAAAA 577

Db	14847	ATGCCATTACCCCTTTGGATCAATTGGAAAGATCAAAGCGAGACGAGATGGTTGATAAGA	14906
Qy	578	TTGCCTTGAAGGACGCCATCAGCAGGCTGAGCGAGCGTGGCCAGCTCATCGTCTACCTGC	637
Db	14907	TCGCCTTAAAGAAGCGATTCTATAACTTAGAGGAAAGAGAGCGGTTATCGTTTATTAC	14966
Qy	638	GCTATTACAAGGATCAGACACAGTCTGAGGTAGCCAGAGCGCTCTAGGGATTCGCAGGTCC	697
Db	14967	GTTATTATAAAGATCAAACCCAAATCCGAGGTGCGCAAAGACTCGGAATCTCCCAAGTAC	15026
Qy	698	AGGTCTCGCGTCTGGAAAAGCGTATCCTGCTAACGATCAAGGAGCAAATTGAACATTAG	756
Db	15027	AAATCGAGACTTGAAAAGAAAATCCTGAAACGATGAAAGCGAATATAGCTGAATAG	15085

RESULT 14  
CP000764\_28/c  
WPCOMMENT

Sequence split into 41 fragments LOCUS CP000764 Accession CP000764

Fragment	Name	Begin	End
CP000764_00		1	110000
CP000764_01		100001	210000
CP000764_02		200001	310000
CP000764_03		300001	410000
CP000764_04		400001	510000
CP000764_05		500001	610000
CP000764_06		600001	710000
CP000764_07		700001	810000
CP000764_08		800001	910000
CP000764_09		900001	1010000
CP000764_10		1000001	1110000
CP000764_11		1100001	1210000
CP000764_12		1200001	1310000
CP000764_13		1300001	1410000
CP000764_14		1400001	1510000
CP000764_15		1500001	1610000
CP000764_16		1600001	1710000
CP000764_17		1700001	1810000
CP000764_18		1800001	1910000
CP000764_19		1900001	2010000
CP000764_20		2000001	2110000
CP000764_21		2100001	2210000
CP000764_22		2200001	2310000
CP000764_23		2300001	2410000
CP000764_24		2400001	2510000
CP000764_25		2500001	2610000
CP000764_26		2600001	2710000
CP000764_27		2700001	2810000
CP000764_28		2800001	2910000
CP000764_29		2900001	3010000
CP000764_30		3000001	3110000
CP000764_31		3100001	3210000
CP000764_32		3200001	3310000
CP000764_33		3300001	3410000
CP000764_34		3400001	3510000
CP000764_35		3500001	3610000
CP000764_36		3600001	3710000
CP000764_37		3700001	3810000
CP000764_38		3800001	3910000
CP000764_39		3900001	4010000
CP000764_40		4000001	4087024

Continuation (29 of 41) of CP000764 from base 2800001 (CP000764 *Bacillus cereus* subsp.

Query Match 45.2% Score 342; DB 14; Length 110000;  
Best Local Similarity 67.6% Pred. No. 6.5e-98;  
Matches 480; Conservative 0; Missmatches 230; Indels 0; Gaps 0;

Qy 47 ACCAAGTGAAGAGATTGATAGCCAAGAGCCAAGCTGGCGATACGGATGCACGTGAGCTTC 106

Db 53779 ACGAGTTAAAAGCGTTAATTCAAAAAAGTCAGATGGAGATCAAGGGCGAGAGATACGA 53720  
 Qy 107 TCGTGAATAGCAATATCAGACTGGCTGGTCCGTCGTCAGGGCTTATCAACCGGGGT 166  
 Db 53719 TTGTTCAAAGTAATATGCGCCTTGTATGGTCAGTTGACAACGTTCTCAATCGAGGAT 53660  
 Qy 167 ATGAAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTCAAGGGCGTTGACAAGT 226  
 Db 53659 ATGAACCCGACGACTTATTCAAATTGGCTGATTGGCTCTTGAATCAGTAGACAAAT 53600  
 Qy 227 TCGATCTTCGTACGATGTGAGATTTGACCTATGCGGTGCCAATGATCGGAGAAA 286  
 Db 53599 TTGATTTATCTTTGATGTGAAGTTTCAACTTACGCAGTCCAATGATTATTGGTGAAA 53540  
 Qy 287 TTCAACGCTTTTGGCGATGACGGTAACGGTAAGGTCACTCGATCGTAAAGAACAG 346  
 Db 53539 TTCAACGTTTTTACGTGATGACGGTCAGTAAAGTGAGCCGATCGTAAAGAACGG 53480  
 Qy 347 CGAATAAGGTGGCGATCAAAGGATGAATTGTACAAGCAATTGGCGTGC0000CACGA 406  
 Db 53479 GTAATAAGATTGAAAGATGAAAGATGAGCTTCAAAAGATACCGAAGGGCACCAACCA 53420  
 Qy 407 TCGCAGAAGTGGCAGAAGCAGTGGGAATCACGCCGGAGGAAGTAGTCTTGCGCAAGAGG 466  
 Db 53419 TTAATGAAGTTGCAAGAGGCGTTGAACTAACACCGAAGAAGTTGTTCTGGCGCAAGAGG 53360  
 Qy 467 CAAGCAGAGGCCCTCCATCCATGAGACCGTTTGAAAATGACGGCGATCCATCA 526  
 Db 53359 CAAGTCGTACGCCATCTTCTATACATGAAACTGTATATGAAATGATGTTGATCCAATCA 53300  
 Qy 527 CACTGATCGATCAGATAGCGGATGAAGGTGTGAACAAAGTGGTTGAGAAAATTGCCCTGA 586  
 Db 53299 CGATTTGGATCAAATTGGCGATCAAGCGAAACGAAATGGTTGATAAGATTGCTTAA 53240  
 Qy 587 AGGACGCCATCAGCAGGCTGAGCGAGCGTGAGCAGCTCATCGTCAACCTGCCATTACA 646  
 Db 53239 AAGAACCAATAAGAGAACTAGATGACCGAGAACGCTTAATTGTATATTGCCATTATA 53180  
 Qy 647 AGGATCAGACACAGTCTGAGGTAGCAGAGCGCTAGGATTGCCAGGTCCAGGTC 706  
 Db 53179 AAGATCAAACCAATCGGAAGTAGCAGAGAGAATTGAAATTCAACAGTCCAAGTTGCA 53120  
 Qy 707 GTCTGGAAAAGCGTATCCTGCTAACGATCAAGGAGCAAATTGAAACATTAG 756  
 Db 53119 GGTTGGAAAAGAAAATTGAAGCAGATGAAAGATCGGATAGATGAGTAG 53070

## RESULT 15

X63757

LOCUS X63757 1920 bp DNA linear BCT 18-APR-2005  
 DEFINITION B. megaterium genes spoI AA, spoI AB, spoI AC, ppb and spoVA.  
 ACCESSION X63757 S46395  
 VERSION X63757.1 GI:39650  
 KEYWORDS ppb gene; spoI AA gene; spoI AB gene; spoI AC gene; spoVA gene.  
 SOURCE Bacillus megaterium  
 ORGANISM Bacillus megaterium  
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 REFERENCE 1  
 AUTHORS Tao, Y. P., Hudspeth, D. S. and Vary, P. S.  
 TITLE Cloning and sequencing of the *Bacillus megaterium* spoI A operon  
 JOURNAL Biotechnol 74 (7-8), 695-704 (1992)  
 PUBMED 1391049  
 REFERENCE 2 (bases 1 to 1920)  
 AUTHORS Vary, P. S.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-JAN-1992) P. S. Vary, Northern Illinois University,  
 Dept. of Biological Sciences, DeKalb, Illinois 60115-2861, USA  
 COMMENT On Jun 23, 2005 this sequence version replaced gi:258183.  
 FEATURES Location/Qualifiers

source 1. . 1920  
 /organism="Bacillus megaterium"  
 /mol\_type="genomic DNA"  
 /strain="PV361"  
 /db\_xref="taxon:1404"  
 /map="fragment E, on 339kb NotI fragment"  
 /clone="pYP4, pYP5, pYP9"  
 gene <1. . 77  
 /gene="pbp"  
 CDS <1. . 77  
 /gene="pbp"  
 /codon\_start=3  
 /transl\_table=11  
 /protein\_id="CAA45285.1"  
 /db\_xref="GI:39651"  
 /db\_xref="UniProtKB/TrEMBL:Q45319"  
 /translation="YRCCFLVEDI QTYI LYFLSNI LNS"  
 misc\_signal 81. . 87  
 /note="OA-box"  
 gene 110. . 141  
 /gene="spolI AA"  
 -35\_signal 110. . 118  
 /gene="spolI AA"  
 -10\_signal 134. . 141  
 /gene="spolI AA"  
 gene 159. . 164  
 /gene="spolI AA and spolI AB"  
 RBS 159. . 164  
 /gene="spolI AA and spolI AB"  
 gene 172. . 522  
 /gene="spolI AA"  
 CDS 172. . 522  
 /gene="spolI AA"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAA45286.1"  
 /db\_xref="GI:580788"  
 /db\_xref="GOA: P35147"  
 /db\_xref="InterPro: IPR002645"  
 /db\_xref="InterPro: IPR003658"  
 /db\_xref="UniProtKB/Swiss-Prot: P35147"  
 /translation="MSLSI ELEFKQDVLLVRLTGELDHHTAEELRSKI TEAI EKESI S  
 HLI LNLQHLTMDSSGLGVI LGRYKQMHNNGEMVCAI SPAI ERLFNMSGLFKI I RF  
 EPNEANALQKLGV/A"  
 gene 522. . 965  
 /gene="spolI AB"  
 CDS 522. . 965  
 /gene="spolI AB"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAA45287.1"  
 /db\_xref="GI:39653"  
 /db\_xref="GOA: P35148"  
 /db\_xref="InterPro: IPR003594"  
 /db\_xref="InterPro: IPR010194"  
 /db\_xref="UniProtKB/Swiss-Prot: P35148"  
 /translation="MKNQMLQFSALSQNESFARVTVASFI TQLDPTMDELTEI KTVW  
 SEAVTNAI I HGYESNSDGVYI SVTLHEDGVVELM RDEGI G DNVDEAKQPLYTTKP  
 DLERSGMGFTI MENFMDEI RVESTVLEGGTTLYLKKHLTNSKALCN"  
 gene join(965. . 969, 978. . 1739)  
 /gene="spolI AC"  
 RBS 965. . 969  
 /gene="spolI AC"  
 CDS 978. . 1739  
 /gene="spolI AC"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAA45288.1"  
 /db\_xref="GI:39654"

```

/ db_xr_ef = " GOA: P35145"
/ db_xr_ef = " InterPr o: I PR000943"
/ db_xr_ef = " InterPr o: I PR007624"
/ db_xr_ef = " InterPr o: I PR007627"
/ db_xr_ef = " InterPr o: I PR007630"
/ db_xr_ef = " InterPr o: I PR013324"
/ db_xr_ef = " InterPr o: I PR013325"
/ db_xr_ef = " UniPr ot KB/Swiss-Pr ot : P35145"
/ transl ation=" MDVEVKKNKNEPYLKDHEVKDLIKRSQQGDQIARDTIVQKNMRL
VWSVVQRFI NRGYEPDDL FQI GCI GLLKSVDKFDLSYDVKFSTYAVPM I GEI QRFI R
DDGTVKVSRSLKEMSNKI RKA KDEL SKLLGRVPTVAEVAEHL DLTPEEVVLAQEANRA
PSSI HETVYENDGDPITLLDQIADHTEAKWFDKIALKEAI EELDEREKLI VYLRYYKD
QTQSEVAARLGI SQVQVSRLEKKI LKQMKLHMNDT"
ter min or
ter min or
mi sc_si gnal
- 35_si gnal
- 10_si gnal
RBS
gene
CDS
1750. . 1767
1772. . 1789
1801. . 1820
/ not e=" 20nt - box"
1829. . 1835
/ not e=" spoVA gene"
1853. . 1860
/ not e=" spoVA gene"
1885. . 1890
/ not e=" spoVA gene"
1899. . >1920
/ gene=" spoVA"
1899. . >1920
/ gene=" spoVA"
/ codon_start=1
/ transl_table=11
/ protein_id="CAA45289. 1"
/ db_xr_ef = " GI : 580789"
/ transl ation=" MENVLYL"

```

## ORIGIN

Query Match 44.9% Score 339.6; DB 14; Length 1920;  
 Best Local Similarity 66.3% Pred. No. 2e- 97;  
 Matches 489; Conservative 0; Missmatches 249; Indels 0; Gaps 0;

Qy	16	AAAAATCCGAGTCAACCAATTCTGACCAATGACCAAGTGAAAGATTGATAGCCAAGAGC	75
Db	996	AAAAATAAAAACGAACCGTATTAAGGATCATGAAGTTAAAGACCTCATCAAGCGCAGT	1055
Qy	76	CAAGCTGGGATAACGGATGCACGTGAGCTTCTCGTGAATAGCAATATCAGACTGGCTGG	135
Db	1056	CAGCAACGTGATCAAATTGCACGAGATACAATOGTCAAAAAAATATGGGGCTGTTGG	1115
Qy	136	TCCGTCTGCCAGGCTTATCAACCGGGGTATGAAGGGATGATTTGTTCAGATCGGT	195
Db	1116	TCGGTTGTTAACGTTATTAAACCGTGGTTATGAGCCCGATGATTATTCAAATTGGA	1175
Qy	196	TGCATTGGCTGCTAACGGCGTTGACAAGTTCGATCTTCGTACGATGTGAGATTTCG	255
Db	1176	TGTATAGGGCTATTAAAGTCCGTTGATAAAATTGACTTATCGTATGATGTTAAATTCA	1235
Qy	256	ACCTATCCGGTGCCAATGATCATCGGAGAAATTCAACGCTTTTGGCGGATGACGGTACG	315
Db	1236	ACATATCCTGTTCAATGATTATTGGTGAATTCAACGATTTCGATGATGGCAGC	1295
Qy	316	GTAAAGGTAGTCGATCGTAAAGAACAGCGAATAAGGTGGGGGATCAAAGGATGAA	375
Db	1296	GTAAAGGTAGCCGTTCTCTGAAAGAAATGAGCAATAAAACGTAAAGCAAAAGACGAG	1355
Qy	376	TTGTACAAGCAATTGGCGGTGCCCCACGATGGCAGAAGTGGCAGAACGAGTGGAAATC	435
Db	1356	CTTCCAAATTACTGGACCGTCCCTACAGTCGCTGAGGTTGCAGAACATTGGATCTC	1415
Qy	436	ACGCCGGAGGAAGTAGTCTTGGCAAGAGGCAAGCAGAGGCGCTTCCCTCCATGAG	495
Db	1416	ACTCCTGAAGAAGTAGTGCTAGCTAAGAAGCTAACGTGCTCCTCGTCTATTACGAG	1475

Qy	496	ACCGTTTTGAAAATGACGGCGATCCCATCACACTGATGATCAGATAGCGGATGAAGGT	555
Db	1476	ACCGTCTATGAAAATGATGGAGATCCAATTACCCCTCTTGACCAAATTGCTGATCATACG	1535
Qy	556	GTGAACAAAGTGGTTGAGAAAATTGCCTTGAAGGACGCCATCAGCAGGCTGACCGAGCGT	615
Db	1536	GAAGCGAAGTGGTTGATAAAATTGCCTTAAAAGAAGCAATTGAAGAACCTGATGAGCGT	1595
Qy	616	GAGCAGCTCATCGTCTACCTGCGCTATTACAAGGATCAGACACAGTCTGAGGTAGCAGAG	675
Db	1596	GAAAAGCTCATTGTCTATTGAGGTATTACAAGGATCAAACCTCAGTCCGAAGTGGGGCA	1655
Qy	676	CGTCTAGGGATTTCGCAGGTCCAGGTCTCGCGTCTGGAAAAGCGTATCCTGCTAACGATC	735
Db	1656	AGGCTGGGAATTTCGCAGGTTCAGGTATCGAGGCTTGAGAAGAAAATCTTAAAACAAATG	1715
Qy	736	AAGGAGCAAATTGAACAT	753
Db	1716	AAGCTGCATATGAATGAT	1733

Search completed: November 4, 2008, 17:36:32  
Job time : 1553 secs

---

SCORE 3.0